

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 9, 2005, 02:08:24 ; Search time 7208 Seconds  
(without alignments)  
13165.125 Million cell updates/sec

Title: US-10-089-612-2

Perfect score: 2493  
Sequence: 1 agcctttccggatgaag.....aaatagcttcacagctaga 2493

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: gb\_est1.\*
  - 2: gb\_est2.\*
  - 3: gb\_hic.\*
  - 4: gb\_est3.\*
  - 5: gb\_est4.\*
  - 6: gb\_est5.\*
  - 7: gb\_est6.\*
  - 8: gb\_gse1.\*
  - 9: gb\_gse2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	713.2	28.6	956	CG165489	PURK0807B
2	626.8	25.1	704	CC431305	PURK0807B
3	620.6	24.9	958	BZ789073	PURK0807B
c 4	539.6	21.6	924	BZ789073	PURK0807B
5	379	15.2	555	BZ407849	OGAC1347M
c 6	355	14.2	578	BZ407849	OGAC1347M
c 7	328.6	13.2	1114	CL984849	ZMWBH000
c 8	328.2	13.2	1062	CL987946	ZMWBH000
c 9	311	12.5	1058	CL986825	ZMWBH000
c 10	292.6	11.7	1035	CL986043	ZMWBH000
11	282.8	11.3	917	CG171826	PURK0807B
12	281.6	11.3	808	CG204991	PURK0807B
13	259.6	10.4	882	CG368356	OGXDR48TV
14	208.2	8.4	788	CG067437	PURK0807B
15	207.8	8.3	582	BZ312667	id27h06.b
c 16	203.8	8.2	760	CG997884	ZUAE557TH
c 17	182	7.3	910	BZ703613	PURK0807B
c 18	180.2	7.2	983	CL254917	ZMWBH000
c 19	175.4	7.0	968	CG139245	PURK0807B
20	170.8	6.9	931	CG158928	PURK0807B
21	170.6	6.8	876	CG127109	PURK0807B
22	170.2	6.8	708	CC433252	PURK0807B
23	168.8	6.8	1025	CG047285	PURK0807B
24	168.4	6.8	905	CC434444	PURK0807B

c 25	155.6	6.2	514	8	BZ327501	BZ327501 id27h06.g
c 26	153.4	6.2	425	9	CC735092	CC735092 OGUG77TV
c 27	153.2	6.1	1038	9	CG420560	CG420560 ZMWBH000
c 28	152.2	6.1	590	8	BZ678438	BZ678438 PUBC677D
c 29	152.2	6.1	609	8	BZ376109	BZ376109 id65403.g
c 30	152.2	6.1	714	8	BZ671566	BZ671566 PUBC677D
c 31	152.2	6.1	791	8	BZ372257	BZ372257 id65403.b
c 32	151.4	6.1	792	8	CC012679	CC012679 PUBH026TD
c 33	146.6	5.9	579	8	BZ372153	BZ372153 id64803.b
c 34	143	5.7	508	8	BH784025	BH784025 fzab01360
c 35	142.8	5.7	779	9	CG069355	CG069355 PUAF91TD
c 36	141.6	5.7	929	9	CG139241	CG139241 PUKA14TD
c 37	141.4	5.7	819	8	CC376588	CC376588 PURK0807B
c 38	140.6	5.6	968	8	CC434443	CC434443 PUNK907B
c 39	139.4	5.6	860	9	CG368347	CG368347 OGXDR48TV
c 40	138.6	5.6	740	8	BZ982826	BZ982826 PUFJY05TD
c 41	138.6	5.6	890	8	CC386965	CC386965 PUBGR47TD
c 42	137.4	5.5	743	8	BZ756164	BZ756164 PUBD019TD
c 43	135.2	5.4	853	8	BZ680379	BZ680379 PUBJ3C94TD
c 44	134.8	5.4	848	8	BZ756167	BZ756167 PUBD019TD
c 45	125.2	5.0	1021	9	CG198973	CG198973 PUPZM68TB

## ALIGNMENTS

RESULT 1  
CG165489  
LOCUS  
DEFINITION  
PURK0807B ZM\_0.6.1.0\_KB\_Zea mays genomic clone ZMWBH0797M15,  
genomic survey sequence.  
ACCESSION  
CG165489  
VERSION  
CG165489.1 GI:34056290  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
1 (bases 1 to 956)  
White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Renick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and  
Bennetzen, J.  
Maize Genomics Consortium  
Unpublished (2003)  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.  
Location/Qualifiers  
1..956  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMWBH0797M15"  
/clone\_lib="ZM\_0.6.1.0\_KB"  
/note="Vector: PCR1-TOPO, Site 1: EcoRI, 0.6-1.0 kb high  
Cot selected genomic DNA library"

Query Match 28.6%; Score 713.2; DB 9; Length 956;  
Best Local Similarity 90.8%; Pred. No. 1.5e-156;  
Matches 872; Conservative 0; Mismatches 73; Indels 15; Gaps 10;

Oy 1397 GGCGATGCTCTTTTATAGGG--AGCACCATTGATTAAAT--GGAAGACATGACACAAG 1453  
DB 3 GGCGTGGTCTCTTTTATAGGGACCAATGATTAAATGGGAGACATGACAGAG 62



Db	301	TCCAATTTCTTAATAATCCCTTGAATCCACATTTCTATTTTGGCCATTTCTTCAAAATTTGGC	360
Qy	2142	ACAAAACTAGGAAAAATTTAATACATTTCTTTGGCATAACATATTTCTAGTGCAAAATGTTAACT	2201
Db	361	ACAAAACTAGGAAAAATTTAATACATTTCTTTGGCATAACATATTTCTAGTGCAAAATGTTAACT	420
Qy	2202	AGATGGCTCAATATTAGCAAACTCTTTTGTAGATTCATTAATTTGCTACATTCGCATA	2261
Db	421	AGATGGCTCAATATTAGCAAACTCTTTTGTAGATTCATTAATTTGCTACATTCGCATA	480
Qy	2262	CTTTTGTAGAAAGTTCAATCAATAATGCTCATTAGCATACTTCATTTTGGAACTTGATTA	2321
Db	481	CTTTTGTAGAAAGTTCAATCAATAATGCTCATTATCATATCATTTTATAGCACTTGATTA	540
Qy	2322	AAACCGCCTTAAATATAGAGCAAGTGACGGATCCA-TTTAAAGGTGATCTCTTAATTTCTT	2380
Db	541	AAACCGCCTTAAATATAGAGCAAGTGACGGATCCA-TTTAAAGGTGATCTCTTAATTTCTT	600
Qy	2381	ACTTCTCATCTTGTGGCTTATGTTTATATATATATATGTTGGTGGTGTGAATGATGTTCTTAC	2440
Db	601	ACTTCTCATCTTGTGGCTTATGTTTATATATATATGTTGGTGGTGTGAATGATGTTCTTAC	660
Qy	2441	ACCACTACACACAGTTGGACATATATATGGAATAGCTTCACAGTCTTAGA	2493
Db	661	ACCACTACACACAGTTGGACATATATGGAATAGCTTCACAGTCTTAGA	713
RESULT 4			
BZ789073/c			
LOCUS			
DEFINITION			
PUPDF97D ZM 0.61.0 KB Zea mays genomic clone ZMBMTa292N14,			
genomic survey sequence.			
ACCESSION			
BZ789073			
VERSION			
BZ789073.1 GI:28983189			
KEYWORDS			
GSS.			
SOURCE			
Zea mays			
ORGANISM			
Zea mays			
REFERENCE			
AUTHORS			
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,			
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and			
Bennetzen,J.			
TITLE			
Maize Genomics Consortium			
JOURNAL			
Unpublished (2003)			
COMMENT			
Other GSSs: PUPDF797B			
Contact: Cathy Whitelaw			
TIGR			
9712 Medical Center Drive, Rockville, MD 20850, USA			
Tel: 301-838-5843			
Fax: 301-838-0208			
Email: whitelaw@cigr.org			
Seq primer: TP			
Class: sheared ends.			
FEATURES			
Source			
1. .924			
Location/Qualifiers			
/organism="Zea mays"			
/mol_type="genomic DNA"			
/strain="B73"			
/db_xref="taxon:4577"			
/clone="ZMBMTa292N14"			
/clone_lib="ZM 0.6 1.0 KB"			
/notes="Vector: pCR1-TOPO; Site 1: ScaRI; 0.6-1.0 kb high			
Cor selected genomic DNA library"			
ORIGIN			
Query Match			
21.6%; Score 539.6; DB 8; Length 924;			
Best Local Similarity			
99.1%; Prd. No. 9.8e-116;			
Matches 553; Conservative 0; Mismatches 4; Indels 1; Gaps 1;			
Qy 1937 TGGTCAGGTTGGGCTCGGCTTTGGTGTAGTATCTTTGTATATAAATAATTTCTCTCTT			
1996			

Db 924 TGTGGAGTTGGGCTCCGCTTTGGTGTATGTTGTATATATAAATAATTTCTCTCTT 865  
Qy 1997 TTGTCACTTCAATATATGACTTAATATTTTATGTAGCAATGCCAATTTTTTATGAGT 2056  
Db 864 TTGTCACTTCAATATATGACTTAATATTTTATGTAGCAATGCCAATTTTTTATGAGT 805  
Qy 2057 TGGCACTATAGCAAAACTATATCCATTTTCTAATAGTCTTTGAAATCCACATTTCTA 2116  
Db 804 TGGCACTATAGCAAAACTATATCCATTTTCTAATAGTCTTTGAAATCCACATTTCTA 745  
Qy 2117 TTTTATGCAATTTCTTCAAAATTTGGCAAAACTAGGAAATTTAATACATTTCTGCCATA 2176  
Db 744 TTTTATGCAATTTCTTCAAAATTTGGCAAAACTAGGAAATTTAATACATTTCTGCCATA 685  
Qy 2177 ACATATTTCTAGTCAAAATTTAACTAGATTTGCTCAATATAGCAAACTCTTTTGTAGA 2236  
Db 684 ACATATTTCTAGTCAAAATTTAACTAGATTTGCTCAATATAGCAAACTCTTTTGTAGA 625  
Qy 2237 TTAATTAATATGCTACATTTGCAATCTTTTATAGAGTTTCAATATAGCTCTATAGC 2296  
Db 624 TTAATTAATATGCTACATTTGCAATCTTTTATAGAGTTTCAATATAGCTCTATAGC 565  
Qy 2297 ATACTTCAATTTTAGCACTTGAATTAACCGCTTAAATAGAGCAAGTACGATCCA 2356  
Db 564 ATACTTCAATTTTAGCACTTGAATTAACCGCTTAAATAGAGCAAGTACGATCCA 505  
Qy 2357 -TTTAAAGTCAATTTCTTAAATTTCTTCTATCTTCTGCTGCTTATATATATGT 2415  
Db 504 CTTTAAAGTCAATTTCTTAAATTTCTTCTATCTTCTGCTGCTTATATATATGT 445  
Qy 2416 GTGGTGGTTCGAATGATTTCTTACACCACTTACACCACTTACACCACTTACACCACT 2475  
Db 444 GTGGTGGTTCGAATGATTTCTTACACCACTTACACCACTTACACCACTTACACCACT 385  
Qy 2476 ATAGCTTACAGTCTAGA 2493  
Db 384 ATAGCTTACAGTCTAGA 367

## RESULT 5

BZ407849 555 bp DNA linear GSS 04-DEC-2002  
LOCUS OGAC134TM\_ZM\_0.7\_1.5\_KB\_Zea\_mays\_genomic\_clone\_ZM8Ma0029E19,  
DEFINITION Genomic survey sequence.

ACCESSION BZ407849

VERSION BZ407849.1 GI:26037521

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
Clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 555)

Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Reinick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other\_GSSs: OGAC134TC

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: Sheared ends.

Location/Qualifiers

1..555

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

## FEATURES

source

1..555

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZM8Ma0029E19"  
/clone\_lib="ZM\_0.7\_1.5\_KB"  
/note="Vector: pBCSK-; Site.1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 15.2%; Score 379; DB 8; Length 555;  
Best Local Similarity 98.1%; Pred. No. 5.2e-78;  
Matches 415; Conservative 0; Mismatches 5; Indels 3; Gaps 3;  
Qy 1 AAGCTTTTCGGTGAATGAAGCACCTGTAAATCTTAACAGCATGCTGAAACAAATAGTTA 60  
Db 135 AAGCTTTTCGGTGAATGAAGCACCTGTAAATCTTAACAGCATGCTGAAACAAATAGTTA 194  
Qy 61 GCTGTGTTTTTGGAGGACCTTCGGAAGATGAAGGCCCAACACATCCCATGATCAAGTC 120  
Db 195 GCTGTGTTTTTGGAGGACCTTCGGAAGATGAAGGCCCAACACATCCCATGATCAAGTC 254  
Qy 121 CCATGACTTGCAGAAAAAGCAATTTTATC-AAAATTTCTCATAAACACTTGAACAAAT 179  
Db 255 CCATGACTTGCAGAAAAAGCAATTTTATC-AAAATTTCTCATAAACACTTGAACAAAT 314  
Qy 180 TTCTCTTTTGAAGAGTGTAGAGCACTAGCAACTGTCTACTATAAAAGGTTCCCAATTTTC 239  
Db 315 TTCTCTTTTGAAGAGTGTAGAGCACTAGCAACTGTCTACTATAAAAGGTTCCCAATTTTC 374  
Qy 240 TGGGTATTAACAAATGCGCATGGTAAATAACAAGAGAAATCTCTACTAAGAGCAGTAAATTTG 299  
Db 375 TGGGTATTAACAAATGCGCATGGTAAATAACAAGAGAAATCTCTACTAAGAGCAGTAAATTTG 434  
Qy 300 GCTAAAAACAATAGTAGAGCAATTTTAAATGTAAATAGGAAATAGGAGCATGCTACTTGTGT 359  
Db 435 GCTAAAAACAATAGTAGAGCA-TTTAAATGTAAATAGGAAATAGGAGCATGCTACTTGTGT 493  
Qy 360 CTTTCAGGTTTTGATGTCTCTCAAAAGTGTGCCCCCTGGGGGAGTGTGCAACACTCAAAA 419  
Db 494 CTTTCAGGTTTTGATGTCTCTCAAAAGTGTG-CCCCCTAGGGGAGTGTGCAACACTCAAAA 552  
Qy 420 TCT 422  
Db 553 TCT 555

## RESULT 6

BZ407846/c

LOCUS BZ407846

DEFINITION OGAC134TC\_ZM\_0.7\_1.5\_KB\_Zea\_mays\_genomic\_clone\_ZM8Ma0029E19,

genomic survey sequence.

ACCESSION BZ407846

VERSION BZ407846.1 GI:26037516

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
Clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 578)

Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Reinick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other\_GSSs: OGAC134TM

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..578

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

```

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBHD0029819"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

## ORIGIN

```

Query Match      14.2%; Score 355; DB 8; Length 578;
Best Local Similarity 98.2%; Pred. No. 2,38-72;
Matches 380; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy 1 AAGCTTTCCGTTGATGAGCAGCCTGTATATCTTAACAGCATGCTGAAACAAATAGTTA 60
Db |||||
Qy 386 AAGCTTTCCGTTGATGAGCAGCCTGTATATCTTAACAGCATGCTGAAACAAATAGTTA 327
Db |||||
Qy 61 GCTGTGTTTTCAGGACCTTCGAGATGAGGCGCCCAACACATCCCATGCATCAAGTC 120
Db |||||
Qy 326 GCTGTGTTTTCAGGACCTTCGAGATGAGGCGCCCAACACATCCCATGCATCAAGTC 267
Db |||||
Qy 131 CCCATGACTTGCAGAAAGCAAAATTTTATC-AAAATTTCTCATAAACACATTGAAACAT 179
Db |||||
Qy 266 CCCATGACTTGCAGAAAGCAAAATTTTATC-AAAATTTCTCATAAACACATTGAAACAT 207
Db |||||
Qy 180 TTCTCTTTTGAAGTGTAGACGACTAGCACTGCTACTAAAGGTTCCCAATTC 239
Db |||||
Qy 206 TTCTCTTTTGAAGTGTAGACGACTAGCACTGCTACTAAAGGTTCCCAATTC 147
Db |||||
Qy 240 TGGGTATAACAATCGCATGCTAAATAACACAAAGGAAATCTACTAAGAGCAGTAATTTG 299
Db |||||
Qy 146 TGGGTATAACAATCGCATGCTAAATAACACAAAGGAAATCTACTAAGAGCAGTAATTTG 87
Db |||||
Qy 300 GCTAAACAAATAGTGAGCAATTTAATGTAATAGGCAATAGGAGTCATCTGTT 359
Db |||||
Qy 86 GCTAAACAAATAGTGAGCA-TTTAATGTAATAGGCAATAGGAGTCATCTGTT 28
Db |||||
Qy 360 CTTTCAGGGTTTTGATGTCCTCAAAAG 386
Db |||||
Qy 27 CTTTCAGGGTTTTGAGTCTCAAAAG 1
Db |||||

```

## RESULT 7

```

CL984849/c
LOCUS      ZMMBHD0002122.f ZMMBHD Zea mays genomic clone ZMMBHD0002122 5',
DEFINITION genomic survey sequence.

```

```

ACCESSION CL984849
VERSION    CL984849.1
KEYWORDS   GSS.
SOURCE     Zea mays

```

## ORGANISM

```

Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1114)

```

## REFERENCE

```

AUTHORS    Ma,J., SanMiguel,P., Liu,R., Haller,K., Soderlund,C. and
            Bennetzen,J.

```

## TITLE

```

ZMMBH sequences

```

## JOURNAL

```

COMMENT    Unpublished (2004)

```

## CONTACT

```

Contact: Jeff Bennetzen
Bennetzen Lab
The University of Georgia
Department of Genetics, C426a Life Sciences Building, Athens, GA
30602, USA

```

```

Tel: 706-542-3698
Fax: 706-583-0972
Email: maize@uga.edu

```

```

Plate: 0002 row: j column: 22

```

```

Class: BAC ends.

```

## FEATURES

```

source     Location/Qualifiers
            1..1114
1099      GTGCAATCATGTTGCTGATCGATGTCAGGGAGCAATTAGGGTTCACGAGTCAGCGATG 1158
229       CCCGATGGTCCATGCACTATCCAGCTAGGGAGGAGACGTTTCGAGGTTGGGTACCTCTG 170
1039      CAAGATTTCTTATCATGTGACCGGGGTACGGGAATGGCGCATGGGGTTTGGTACTTTCTG 1098
289       GGGGATGGTGGTGGCGGTGAGGGGGCGTTTCACGGGGAAGAGCATGCTTCGATTTAGCAAT 230
979       GGGGCAAGATCGATGGTGGCGATGGTTCGTGATGGAGGGGAACCTTCGCTGAGCAATT 1038
349       AAGCCACTAGGGCTAGCTGCAAGCATGCTAGGCACCACCTCGGGGGCAAGGAGCATGCGA 290
919       AAGCCACTGTGGCTCATCGATGATGATGAGCACAATATCAAGGTGCGAGGTCACACAGA 978
799       ATTAATATGGTGGATGTGCAAGCGACCGCACACGATGGAGGAATGGCTTCACGACGGTGG 858
468       AATAATATGGATCA-CTATGTGGCGCGCGCACACGCTGTAGGAGACCTCCCGGGGGTAG 410
859       GCTACTAGAGCTGGCTAGCTCAACCAATGGAGGGCTCGGTCAAGGTCAAAATTTGTTGCC 918
409       GTCAGTGGTGGAGGCGCATGTGCGGCAACAGAGAGCTGGGCTAGGGCTCAACCTGTTGCC 350
919       AAGCCACTGTGGCTCATCGATGATGATGAGCACAATATCAAGGTGCGAGGTCACACAGA 978
349       AAGCCACTAGGGCTAGCTGCAAGCATGCTAGGCACCACCTCGGGGGCAAGGAGCATGCGA 290
979       GGGGCAAGATCGATGGTGGCGATGGTTCGTGATGGAGGGGAACCTTCGCTGAGCAATT 1038
289       GGGGATGGTGGTGGCGGTGAGGGGGCGTTTCACGGGGAAGAGCATGCTTCGATTTAGCAAT 230
1039      CAAGATTTCTTATCATGTGACCGGGGTACGGGAATGGCGCATGGGGTTTGGTACTTTCTG 1098
229       CCCGATGGTCCATGCACTATCCAGCTAGGGAGGAGACGTTTCGAGGTTGGGTACCTCTG 170
1099      GTGCAATCATGTTGCTGATCGATGTCAGGGAGCAATTAGGGTTCACGAGTCAGCGATG 1158

```

```

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBHD0002122"
/tissue_type="immature ear"
/dev_stage="6-8 weeks"
/lab_host="DH10B"
/clone_lib="ZMMBHD"
/notes="Vector: TOPOcr4; Site_1: EcoRI; Site_2: EcoRI"

```

## ORIGIN

```

Query Match      13.2%; Score 328.6; DB 9; Length 1114;
Best Local Similarity 64.8%; Pred. No. 4,28-66;
Matches 627; Conservative 0; Mismatches 329; Indels 11; Gaps 9;

Qy 266 ACACAAAGGAAATCTCTAAGAGCAGTAATTTGGCTTAAACAAATAGTGAGCAATTTAAT 325
Db |||||
Qy 1005 ACCCTTAAGTAGTACTAATAGGAGTAGTAATTTGGCCAAACAAATATGACAT-TTTAT 947
Db |||||
Qy 326 GTAATAGGGAATAGGAGCATGCAATATCTTGTGTTCTTTCAAGGTTTTGATGTCCTCAAA 385
Db |||||
Qy 946 GCAATGGGTGATACGAGGATAGTAATGTTTCTTCTCGGGTTTCGAGGTCTCTCAAGA 887
Db |||||
Qy 386 GTGTGCCCCCTGGGGCAGTTGCAACACTCAAAATCTACTCGTATACATA-AAGAACAT 444
Db |||||
Qy 886 GTGGGGCCCTC-GAGTAGTTGCAACATATGATCTACTGTTCTCATATATTAACAT 828
Db |||||
Qy 445 GGGCAGCAAAATAGAAACAAATCTCAATTTATGAAAAA---AGGTTCAAATGCTCTATAA 501
Db |||||
Qy 827 GGTCAAAAATAGAAACAAATATATCAACATAGAAACATGGTTTCAGAAATGCTCTATAA 768
Db |||||
Qy 502 TTATGTAGACATTT-TAGAAATTTATTTAGACCAAAACCAATTTAAATTTGTTTAAAT 560
Db |||||
Qy 767 TTATCTATGATTTGTAGAAATTTATTTGAGACCAAAATCTTTAATTTAGTTAAT 708
Db |||||
Qy 561 GAGTTAGATTAATTAATTTATTCAGTTTATAGTTATTTGGACATTTATTTACTTAATA 620
Db |||||
Qy 707 GAATTAGATATGAATTTTAAATCAATTTTACAAACATCAAGG-CATATATTCATTTATCTA 649
Db |||||
Qy 621 TAACCTTAGGGT-TTTAAAGTAATTTTGGTCCCTAGTTGGAGCTAGCTCAGATTGC 679
Db |||||
Qy 648 TAACCTTCAGGATCTTAAAGTATATTTTGGTCCCTAGTTATATCTGTTGTTGTTACTGC 589
Db |||||
Qy 680 TGGTTGATTTCCATAAAGTCGAGTTCTTTTAGCAAAA-ATCCAGGTGAAACAGGGGG 738
Db |||||
Qy 588 GAGTTTATTTCTGTAATAATTGAGGAGCTCTTAAGCAAAATGTCACGGTCAGCTCATACT 529
Db |||||
Qy 739 AGATAGGTGTCACCGATATCTCTAATTTTGTATCGTTGGACGGCAGCATGATGTTCTCAG 798
Db |||||
Qy 528 GACCAAGTGTATACCAATATTCACGAAATATACGCCATATGCGATCACACGGGTGCGACTCG 469
Db |||||
Qy 799 ATTAATATGGTGGATGTGCAAGCGACCGCACACGATGGAGGAATGGCTTCACGACGGTGG 858
Db |||||
Qy 468 AATAATATGGATCA-CTATGTGGCGCGCGCACACGCTGTAGGAGACCTCCCGGGGGTAG 410
Db |||||
Qy 859 GCTACTAGAGCTGGCTAGCTCAACCAATGGAGGGCTCGGTCAAGGTCAAAATTTGTTGCC 918
Db |||||
Qy 409 GTCAGTGGTGGAGGCGCATGTGCGGCAACAGAGAGCTGGGCTAGGGCTCAACCTGTTGCC 350
Db |||||
Qy 919 AAGCCACTGTGGCTCATCGATGATGATGAGCACAATATCAAGGTGCGAGGTCACACAGA 978
Db |||||
Qy 349 AAGCCACTAGGGCTAGCTGCAAGCATGCTAGGCACCACCTCGGGGGCAAGGAGCATGCGA 290
Db |||||
Qy 979 GGGGCAAGATCGATGGTGGCGATGGTTCGTGATGGAGGGGAACCTTCGCTGAGCAATT 1038
Db |||||
Qy 289 GGGGATGGTGGTGGCGGTGAGGGGGCGTTTCACGGGGAAGAGCATGCTTCGATTTAGCAAT 230
Db |||||
Qy 1039 CAAGATTTCTTATCATGTGACCGGGGTACGGGAATGGCGCATGGGGTTTGGTACTTTCTG 1098
Db |||||
Qy 229 CCCGATGGTCCATGCACTATCCAGCTAGGGAGGAGACGTTTCGAGGTTGGGTACCTCTG 170
Db |||||
Qy 1099 GTGCAATCATGTTGCTGATCGATGTCAGGGAGCAATTAGGGTTCACGAGTCAGCGATG 1158
Db |||||

```

Db 169 GCGTAGCTCTAGGTGCTGAGTGGAGTGGAGGAGCTCTCTGGGCTCAAAATCAACAACA 110  
 QY 1159 AGGGCAGTGGGAGCTCTGTCACCATGTTTCGATCAACTAGGAGCATAGAGCTCTAT 1218  
 Db 109 GTGAGGCGCATGGGCTTGGATTTGTTGCTCTAGTGTCTGAGGCTCGATAGAGCTCTAG 50  
 QY 1219 GAAGTTT 1225  
 Db 49 GGAAGTT 43

RESULT 8  
 CL987946/c  
 LOCUS ZM98He0004m18.1 ZM98He Zea mays genomic clone ZM98He0004m18 3',  
 DEFINITION genomic survey sequence.

ACCESSION CL987946.1 GI:52556024  
 VERSION CL987946.1  
 KEYWORDS GSS.

SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 Clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 1062)  
 AUTHORS Ma, J., SanMiguel, P., Liu, R., Haller, K., Soderlund, C. and Bennetzen, J.

TITLE ZM98 sequences  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Jeff Bennetzen  
 Bennetzen Lab  
 The University of Georgia  
 Department of Genetics, C426a Life Sciences Building, Athens, GA  
 30602, USA

FEATURES  
 source  
 1..1062  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZM98He0004m18"  
 /tissue\_type="immature ear"  
 /dev\_stage="6-8 weeks"  
 /lab\_host="DH108"  
 /clone\_lib="ZM98He"  
 /note="Vector: TOPOpcr4; Site\_1: EcoRI; Site\_2: EcoRI"

## ORIGIN

Query Match 13.2%; Score 328.2; DB 9; Length 1062;  
 Best Local Similarity 64.2%; Pred. No. 5.1e-56;  
 Matches 636; Conservative 0; Mismatches 343; Indels 12; Gaps 9;  
 QY 242 GGTATACATCGCATCGTGTAAATACACAAAGGAATCTCTACTAGAGCAGTAAATTTGGC 301  
 Db 1028 GGTATACATCGCATCGTGTAAATACACACCCCTTAAGGTATCTAATAGGAGTAGTAAATTTGGC 969  
 QY 302 TAAACCAATAGTGAGCATTTTATGTAAATAGGAGTATAGGAGCATGCAATCTGTGTTCT 361  
 Db 968 CAACCAATATGACACA--TTATGCAATGGGTATACAGGATAGATGTTTCTTCT 911  
 QY 362 TTCAGGTTTTCATGCTCTCAAAAGTGTGCCCTCGGGGAGTGTGCAACACTCAAAATC 421  
 Db 910 TCTCGGTTTCAGGCTCTCAAGAGTGGGGCCCTC-GAGGTAGTTGCAACATTATGAATC 852  
 QY 422 TACTCTATACATA-AAGAACATGGGCACAAATAGAAACAATACTCAAAATTATGAA 480  
 Db 851 TACTCTGTGTTTATATAAATCATGTCTCAAAATATAGAAACAATATATCAACATGAA 792

QY 481 A---AGGTTCAATGGTCTCTATTAATTAATTTAGACATTT--TAGAATTTATTTAGACCA 536  
 Db 791 ACATGGTTCAGATGGTCTCTATTAATTAATTTCTATGTAATTTGAGAAATTTATTTGAGACCA 732  
 QY 537 AACCAATTTAAATTTGGTTTAAAAATGAGTTAGATATTAATTAATTTATTCAGTTATAGTTAT 596  
 Db 731 AATCATTTAAATTTAGATTTAAAAATGAATATAGATATGAATTTTAAATCATTTTACACAAAT 672  
 QY 597 TTGGGACATTTATTTTACTTAATACTTAATCTTAGGGT--TTTAAAAAGTAAATTTTGGTCC 655  
 Db 671 CAAGG-CATATATTCATTTTCTATTAATCTTCAAGGTATCTAAGATATATTTTGGTCC 613  
 QY 656 CTAGTTGGAATAGCTAGATTCGTTGATTTTCATTAATAAAGTCGAGGTTCTCTTTAGCA 715  
 Db 612 CTAGTTATAAATCTGTTGCTGCTGCGAGTTTATTTCTGTAAAAATTTGAGAGCTCTTTAAGCA 553  
 QY 716 AAA-ATCCACCGGTGAACAAGGGGGAGATAGTGTGTGACCGATATCTCTAAATTTTGTATCG 774  
 Db 552 AATGGTCAAGGCTCAGCTCATACTGACCAAGTGTCTAACCAATATCCAGATATCAGCA 493  
 QY 775 TTGAGCGGCACATGATGTTCTCAGATTAATGTTGATGTGCAAGCGACGCGACACAGAT 834  
 Db 492 TTGATACACAGGCTGCGATCTCGAATAAATGGATCA-CTATGTGGGCGCGCACACGGT 434  
 QY 835 GGAGGAATGGCTTCACAGCGGTGGGCTACTAGAGCTGGCTACGTCAACCAATGGAGGGCT 894  
 Db 433 GGTAGGAGACCTTCGCGGGGGTGGTCTCTGTGTAGAGCCATGTGCGCAACAGAGAGCT 374  
 QY 895 CGGTCAAGGTCAAAATTTTGTCCAAAGCCTCTGTGCTCAGATGATGATGATGATGATGATG 954  
 Db 373 CGGTAGGGCTACAACTGTGTCCAAAGCCTTAGGGCTACGTCAAGCATGCTTAGGACCA 314  
 QY 955 TATCAAGGTTCAGAGGCTCAACACAGAGGGGCAAGATCGATGTCAGTGTCTCTCGATGG 1014  
 Db 313 CCTCGGGGCAAGGGAAGCTGCGAGGGGATGGTGGCGGTGAGGGGCGCTTCAAGGCA 254  
 QY 1015 AAGGGGAAATCTTCGGTGAAGCAATTCAGATTTCTATCATGTGACCGGGTCAGGGAATGG 1074  
 Db 253 AAGAGCATGCTTCGATTTAGCAATTCCTGATGCTTCCATGCTATCCAGCTAGGGAAGGG 194  
 QY 1075 GGCATAGGGTTTGGTACCTTCTGTCGATCATCATGTCGTGTATTCATGTCGATGTCGATG 1134  
 Db 193 ACCTTCAGGTTTGGGTACCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 134  
 QY 1135 ATTAGGTTTACAGTTCAGGATGACGGGATGACGGGATGTCGGGCTTTGTGTACCATGTTCCAT 1194  
 Db 133 TCCTGTGGGCTCAAAATCAACACAGTGAAGCGGATGGGTTGGATTTGTTGTGCTTAGG 74  
 QY 1195 CAATAGGAGCATAGAGCTCTATCAAGTTT 1225  
 Db 73 TCCTAGGCTCGATAGAGCTCTAGGGAAGTT 43

RESULT 9  
 CL996825/c  
 LOCUS ZM98He0009105.1 ZM98He Zea mays genomic clone ZM98He0009105 3',  
 DEFINITION genomic survey sequence.  
 ACCESSION CL996825  
 VERSION CL996825.1 GI:52566923  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 Clade; Panicoideae; Andropogoneae; Zea.  
 REFERENCE 1 (bases 1 to 1058)  
 AUTHORS Ma, J., SanMiguel, P., Liu, R., Haller, K., Soderlund, C. and Bennetzen, J.  
 TITLE ZM98 sequences  
 JOURNAL Unpublished (2004)  
 COMMENT Contact: Jeff Bennetzen

Bennetzen Lab.  
The University of Georgia  
Department of Genetics, C426a Life Sciences Building, Athens, GA  
30602, USA  
Tel: 706-542-3698  
Fax: 706-583-0972  
Email: maize@uga.edu  
Plate: 0009 row: 1 column: 05  
Class: BAC ends.

Location/Qualifiers  
1. .1058  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBF0009105"  
/tissue\_type="immature ear"  
/dev\_stage="6-8 weeks"  
/lab\_host="DH10B"  
/clone\_lib="ZMMBF"  
/note="Vector: TOPOper4; Site\_1: EcoRI; Site\_2: EcoRI"

# FEATURES

source

# ORIGIN

Query Match 12.5%; Score 311; DB 9; Length 1058;  
Best Local Similarity 64.4%; Pred. No. 5.7e-62;  
Matches 638; Conservative 0; Mismatches 340; Indels 13; Gaps 11;

Qy 242 GGTATAACATCGCATGTAATAACACACAGGAATCCTACTAGACGAGTAATTTGGC 301  
Db 1077 GGTTTAGAAAGTAACCAACACACACCTTTAAGGTAAATCGTA-TAGGAGTAGTAATTTGGC 969

Qy 302 TAAACATATAGTACGATTTTAATGTATAGGAATAGGAGCATGCAATCTGTGTCT 361  
Db 968 -CAACATATAGCA-TTTATGCAATGGGTGATACGAGGATAGATGTTTCTCTCT 911

Qy 362 TTCGGGTTTTCATGCTCTCAAAAGTGTGCCCTCGGGGAGTTGGCACTCAAAATC 421  
Db 910 TCTCGGTTTCAGAGTCTCAAGAGTGGGGCCCTC-GAGGTAGTTGCAACATATATGAATC 852

Qy 422 TACTGTATACATA-AAGAAACATGGGACAAATAAGAAACAACTCAATTTATGAA 480  
Db 851 TACTGTGTTTCATATATAACATGTCACAAATAGAAACATATATCAACATGAA 792

Qy 481 A---AGTTCAATGCTCTATTAATTTATGTAGACATTT-TAGAAATTTATTTAGACAA 536  
Db 791 ACATGTTCAAGATGCTCTATAATTTATCTATGTTATTTGTAGAAATTTATTTGAGACAA 732

Qy 537 AACCATTTAAATTTGGTTTAAATGATGTAGATTAATATTTATTCAGTTTATAGTTAT 596  
Db 731 AATCATTTAATTTAGATTTAAATGATGATGATGATGATTTTATTTATTTATTTACCAAT 672

Qy 597 TTGGACATTTATTTACTTAATTAATCTTCTAGGGT-TTAAAGTAATTTTGGGTCC 655  
Db 671 CAAGC-CATATATTCATTTATCTAATCTTCAAGGTATCTAAAGTATATTTTGGTCC 613

Qy 656 CTAGTTGGAACCTAGCTCAGATGCTGTTGATTTCCATAAAGTCGAGTCTCTTTAGCA 715  
Db 612 CTAGTTAATACCTGGTGTACTCGGAGTTTATTTCTGTAATTTAGAGGCTCTTAGCA 553

Qy 716 AAA-ATCCAAGGTGAACAAAGGGGAGATAGGTGTGACCGATATCTCTAAATTTTATCG 774  
Db 552 AAATGTCACGGTCACTCATCTGACCAAGTGTCTAACCAATATTCAGCAATATACGCA 493

Qy 775 TTGAGCGGCATGATGCTCTCAGATTAAATGGTGGATGTCAGCGGCGGCACGAT 834  
Db 492 TTGATACACCGGTGCGATCTCGAATAATGATGATCA-CTATGTTGGCGGCGCACGGT 434

Qy 835 GGAGGAATGGCTTTCACGACGGTGGCTACTAGAGCTGGCTACGTCAACCAATAGAGGGCT 894  
Db 433 GGTAGGAGACCTCCCGGGGTAGGTCACTGCTGAGAGCCATGTCGGGCAACAGAGAGCT 374

Qy 895 CGGTCAAGTCAAAATTTGTCGAAGCCACTGTGGCTCACGATGATGCGATTGAGCACA 954

Db 373 CGGCTAGGGCTACAACTGTTGCCAAGCCACTAGGGCTAGCTCAAGCATGCTAGGCACCA 314  
Qy 955 TATCAAGCTCGAGGCTCAACACAGAGGGCGAAGATCGATGCTGCTGCTGCTGATGG 1014  
Db 333 CCTCGGGGCAAGGACAGCAGCTGAGAGGGGCAATGTTGGCGGCTTCAAGCGCA 254  
Qy 1015 AAGGGGAAACTTTCGGTGAAGCAATTCCTATCATGTGACCGGGTCAGGGAAATGG 1074  
Db 253 AAGAGCATGCTTCGATTAGCAATTCGCGCATGCTCCATGCACTATCAGCTAGGGAAGGG 194  
Qy 1075 GGCATCGGGTTTGGTACCTTCTGCTGACATCATGCTGCTGATCGATGTCAGGGAGC 1134  
Db 193 ACGTTGAGGTTGGGTACCTGCTGGCGTAGCTCTAGGTGCTGAGTGGAAATCGAGGGAGC 134  
Qy 1135 ATTAGGTTTCAGAGTCAGCGATGACGGGCGATGGGGAGCTTGTGTACCATGCTTGGAT 1194  
Db 133 TCTGTGGGCTCAATCAACAGCATGAGGCGCATGGGTTGGATTGTTGTGCTTAGG 74  
Qy 1195 CAACCTAGGAGCATAGAGCTCTATGAAGTTT 1225  
Db 73 TCGTCAAGGTCGATAGAGCTCTAGGGAAGTT 43

RESULT 10  
CL996043/c  
LOCUS  
DEFINITION  
CL996043  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CL996043 1035 bp DNA linear GSS 23-SEP-2004  
ZMMBF0008k11.f ZMMBF Zea mays genomic clone ZMMBF0008k11 5',  
genomic survey sequence.

CL996043  
CL996043.1 GI:52565370  
GSS.

Zea mays  
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 1035)  
Ma.J., SanMiguel,P., Liu,R., Haller,K., Soderlund,C. and  
Bennetzen,J.

TITLE  
JOURNAL  
COMMENT

ZMMBF sequences  
Unpublished (2004)  
Contact: Jeff Bennetzen  
Bennetzen Lab  
The University of Georgia  
Department of Genetics, C426a Life Sciences Building, Athens, GA  
30602, USA  
Tel: 706-542-3698  
Fax: 706-583-0972  
Email: maize@uga.edu  
Plate: 0008 row: k column: 11  
Class: BAC ends.

FEATURES  
source

1. .1035  
Location/Qualifiers  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBF0008k11"  
/tissue\_type="immature ear"  
/dev\_stage="6-8 weeks"  
/lab\_host="DH10B"  
/clone\_lib="ZMMBF"  
/note="Vector: TOPOper4; Site\_1: EcoRI; Site\_2: EcoRI"

ORIGIN

Query Match 11.7%; Score 292.6; DB 9; Length 1035;  
Best Local Similarity 63.3%; Pred. No. 1.2e-57;  
Matches 627; Conservative 0; Mismatches 349; Indels 15; Gaps 11;

Qy 242 GGTATAACATCGCATGTAATAACACAGGAATCCTACTAGACGAGTAATTTGGC 301  
Db 1025 GGTTAGAAGCTAAACCAACACACCCCTTTAAGTATCTATAGGAGTAGTAATTTGCC 966



302 TAAACCAATAGTGTAGCAATTTTAAATGTAATAGGGAATAGAGGATGCAATACACTGTGTCT 361  
 965 --AAACCAATATGAACA-TTAAATGCAATGGGTGATAGAGGATAGAAAT--GTTTTCTCT 911  
 362 TTCAAGGTTTGTATGTCTCAAAAGTGTGCCCCCTGGGCGCATGTGCAACACTCAAAATC 421  
 910 TCTCGGGTTTCGAGTGTGCAAGAGTGGGGCCCTC-GAGGTAGTTGCAACATATGATATC 852  
 422 TACTGTATATACATA-BAGAAACATGGGCAAAAATAGAAACAATCTCAAAATATGAAA 480  
 851 TACTCGTGTTCATATATTAACATAGGTGCAAAAATAGAAACAATATATCAACATGAAA 792  
 481 A---AGGTTCAATAGTCTCTATAATTTATTTAGACATTT-TAGAAATTTATTTAGACAA 536  
 791 ACATGGTTTCAGATGGTCTCTATAATTTTCTATGTATTGTAGAAATTTATTTAGACAA 732  
 537 ACCATTTAAATTTGGTTTAAATAGGTAGATATTTATTTATTTATTTATTTATTTATTT 596  
 731 ATCATTTTAAATTTAGATTTAAATAGGATTTAGATATTTATTTATTTATTTATTTATTT 672  
 597 TTGGGACATTTATTTACTTAACATATACTTCTAGGT-TTTAAAGTAAATTTTGGTCC 655  
 671 CAAGG-CATATATTTATTTATCTATAACTTCCAAGGTATCTAAAAGTATATTTTGGTCC 613  
 656 CTAGTTGCAACTAGCTCAGATTTGCTGTGTTGATTTTCCATAAAAGTCGAGGTCTCTTAGCA 715  
 612 CTAGTTTAACTGTGTGTCTGCGATTTATTTCTGTAAATTTGAGAGCTCTTTAGCA 553  
 716 AAA-ATCCA CGGTGAACAAAGGGGAGATAGGTGTGACCGATATCTCTAAATTTTGTATCG 774  
 552 AAATGGTCAAGCTCAGCTCATATCTGACCAAGTGTCTAAACCAATATCTCCGAATATCAGCA 493  
 775 TTGAGGCGCACATGATGTCTCAGATTAATATGTTGATGTGCAAGCGCGCACACCAT 834  
 492 TTGATATCACCGGTGCGACTCGGAATTAATGATATCTATGTGTGGGGCGGACACCGT 434  
 835 GAGGAATGGCTTCAAGCGGTGGGTCTAGAGCTGGGTCTAGCTCAACCAATGAGGGGT 894  
 433 GGTAGGAGACCTCCGCGGGGTAGGTCTCATGTGTGAGGCAATGTGGGCAACAGAGAGCT 374  
 895 CGGTCAAGGTCAAAATTTGTGCAAGCGCACTGTGCTCAGCATGATGCTGATGAGCA 954  
 373 CGGTAGGGCTTCAACCTGTGCAAGCGCACTAGGCTAGCTCAAGCATGCTAGGCA 314  
 955 TATCAAGGTTCGAGGTCAACCAAGGGGCAAGATCGAGGTGCTAGTGTGCTCGATGG 1014  
 313 CTTGGGGGCAAGGCAAGCTGAGGGGCGATGTGGGCTGAGGGGCGTTCAGGGCA 254  
 1015 AAGGGGAACCTTCGTCGCGCAATTCAGATTTCTATCATGTGACCGGGTCAGGGAATGG 1074  
 253 AAGCATGCTTTCATTTAGCAATTTCCGATGGTCCATGCTCACTATCCGCTAGGAGGG 194  
 1075 GCGCATGGGTTTGGTACCTTCTGTGCAATCATGTGCTGTATCGATGCTCAAGGAGC 1134  
 193 AGCTTCGAGGTGGGTACCTGCTGCGGTAGCTCTAGGTGCTGAGTGAAGTTCGAGGAGC 134  
 1135 ATTAGGGTTCAAGGTCAGGATGAGGGGCACTGTGGGACTTTGTGTCACCATGTTTCAT 1194  
 133 TCTGTGGGCTCAAAATCAACAGTGAAGGCGCATGGGTTGATTTGTTGTGCTAGG 74  
 1195 CAATAGGAGCATAGAGCTCTATGAATTT 1225  
 73 TCGTCAGGTCGATAGAGCTCTAGGGAATTT 43

RESULT 11  
 CG171826  
 LOCUS CG171826 917 bp DNA linear GSS 21-AUG-2003  
 DEFINITION PUUFF31TB ZM\_0.6.1.0 KB zea mays genomic clone ZM0657E14,  
 genomic survey sequence.  
 ACCESSION CG171826  
 VERSION CG171826.1 GI:34062624  
 KEYWORDS GSS.

SOURCE Zea mays  
 ORGANISM Zea mays  
 REFERENCE 1 (bases 1 to 917)  
 AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennett, J.  
 TITLE Maize Genomics Consortium  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Cathy Whitelaw  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-638-5843  
 Fax: 301-638-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TR  
 Class: sheared ends.  
 FEATURES  
 source  
 1..917  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZM0657E14"  
 /clone\_lib="ZM\_0.6.1.0 KB"  
 /note="Vector: PCR4-TOPO, Site\_1: EcoRI; 0.6-1.0 kb high  
 Cor selected genomic DNA library"  
 ORIGIN  
 Query Match 11.3%; Score 282.6; DB 9; Length 917;  
 Best Local Similarity 70.3%; Pred. No. 2.4e-55;  
 Matches 533; Conservative 0; Mismatches 212; Indels 13; Gaps 11;  
 57 GTTAGCTGTGTTTGGAGACCTTCGGAAGATGAAGGCCCCCAACACATCCCATGATCA 116  
 128 GGTGCGTGGGTGAGCTGGGCGCATAGGGTAAGGAACCTTGCCCAATCTTAACCTCATGATCA 187  
 117 AGTCCCCCATGACTTGCACAAAAGGCAAAATTTTATCAAAATTTCTCATAAAACACTTGA 176  
 188 TGTCCCATGTTTGTCTTAAAC-CATCTTTCATAAATAAAGCTCATAGAACATTTGGAAA 246  
 177 CATTTCTCTTTTGAAGGTGAGGACTAGCACTGTCTACTAAAAGGTTCCCAAT 236  
 247 CATTTCTC--TTTAAACATGCGAGGACCTAGCACTTTGACTAC-AAAAAGGTTCCACGT 303  
 237 TCTGGGTATACCAATCGCATGTGTAAATAACACAAAGGAAATCTTCTAAGAGCAGTAAT 296  
 304 TCCAGGTTTAAACAAGGACAAGGTAAATAACACATAGATAATCTTATTAGGAGAAATACT 363  
 297 TTGGCTAAAACAATAGTGAGCAATTTTATGTATAGGAATAGGAGCATGCATCTCTGT 356  
 364 TTCAACAAAACAATAATGAACA-TTTAATGTGTGGGAATAGGGAATAGGAATCTTCTG 422  
 357 GTTCTTTCAGGGTTTTCATGCTCTCAAAAGTGTGCCCCCTGGGCGAGTTCGAACATCA 416  
 423 CTCTCTTGGGGTTTCGAGGTCTCTCAAGGT-AGAGTCTCTCAGGACACATACAAATTTCTC 481  
 417 AATCTACTCTGTATACATA-AAGAAACATGGGCAAAAATAGAAACAATCTCAATTTA 475  
 482 AGATATATCTACATACATACACAAACATGGGAACAAAATAAGAAACATACACAAACA 541  
 476 TGA AAAA--GGTTCAAAATGGTCTATTAATTTATGTAGACATTTTATGAAATTTATTTAGAC 533  
 542 TGA AAAACATGTTTAAATGGTCTATTAATTTATAGACA-TTTTATAATTTATTTAATAC 600  
 534 CA AAACCAATTAATTTGTTTAAATTCAGTTTAGATTTAATTTATTTACAGTTTATAGT 593  
 601 CA AAACCAATTTACTTTTGGATTTAAATGAATGAATAGATATGACTTTTATGCTATTTA-AGT 559  
 594 TATTTGGGACATTTTATTTACTTAACTATAACTTCTAGGGTTTAAAGTAAATTTTGGGT 653





```

/clone="ZMBR040652H24"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      10.4%; Score 259.6; DB 9; Length 882;
Best Local Similarity 67.5%; Pred. No. 6.7e-50;
Matches 498; Conservative 0; Mismatches 229; Indels 11; Gaps 9;

Qy 242 GGTATACATCGCATGGTAAATAACAACAAGGAATCTCTACTAAGACGAGTAATTGGC 301
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 149 GGTATTAGAAAGCTAAACAACAACACACCCTAAGGTAATCTCTAATAGGAGTAGTAATTGGC 208
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 302 TAAAAACAATAGTGAGCATTTTAAATGTAATAGGGAATAGGAGCATGCAATCTTGCTCT 361
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 209 CAAACAATTAATGAACA-TTTAATGCATGGGTGATACAGGATGAGATGTTTCTCTCT 267
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 362 TTCAGGTTTTCATGCTCTCAAAAGTGTGCCCTCTGGGCGAGTTGCAACACTCAAAATC 421
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 268 TCTCGGTTTCGAGGTCTCAAGAGTGGGGCCCTC-GAGGTAGTTGCAACATTAACAATC 326
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 422 TACTCGGTATACATA-AAGAAACATGGGCAAAATAAGAAACAATACTCAAAATTATGAA 480
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 327 TACTCGGTTCATATAATAAACAATGTCACAAATAGAAACAATATCAACATGAA 386
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 481 A---AGGTTCAAAATGGTCTCTAATTTATTTAGACATTT-TAGAAATTTATTTAGACAA 536
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 387 ACATGGTTGAGAAATGGTCTAATTTATTTCTATATTTTGTAGAAATTTATTTAGACAA 446
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 537 AACCAATTAATTTGGTTTAAATAGATGATATTAATTTATTCAGTTTATAGTTAT 596
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 447 AATCAATTAATTTAGATTTAAATGAATAGATATGAATTTTAAATCATTTTACAACAT 506
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 597 TTGGGCAATTTTACTTAATCAATTAATCTAGGGT-TTTAAAGTAAATTTTGGGTCC 655
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 507 CAGG-CATATATTCATTTATCTTAACCTTCAGAGTCTCAAAAGTATATTTTGGTCC 565
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 656 CTAGTTGGAACTAGCTCAGATTCCTGGTTGATTTTCCATAAAAGTCGAGGTTCCTTAGCA 715
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 566 CTAGTTAATACTGGTGTGTACTCGGAGTTTATTTCTGTAATTTGAGGAGCTCTTAAGCA 625
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 716 AAA-ATCCAGGTGAACAAGGGGAGATAGGTGTGACCGATATCTCTAAATTTGATCG 774
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 626 AATGTGTACGGTCAAGTCTATCTGACCAAGTGTACCAATATCCAGAAATATCAGCA 685
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 775 TTGACGGCACAATGATGTCTCAGATTTAAATGGTGAATGTGCAAGCGACGCGACACGAT 834
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 686 TTGATCACAACGGGTGGGACTCGGAATAAATGGATCA-CTATGTGGGGGGCGCACGGT 744
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 835 GGAGGAATGGTCTCAACGAGTGGGCTACTAGAGCTGGCTACGTCACCAAGTGGAGGCT 894
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 745 GGTAGGAGACCTCCCGGGGGTAGGTCACTGTGAGAGCCATGTCCGCGCAACAGAGACT 804
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 895 CGGTCAAGTCAAAATTTGTCCAAAGCCACTGTGGCTCAAGTGAAGTTCGATTTAGACACA 954
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 805 CGGTAGGGCTACAACTGTGCGCAAGCCACTAGGGCTAGCTCAAGCATCTAGGCACCA 864
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 955 TATCAAGGTCCAGGGTCA 972
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 865 CCTCGGGGGCAAGGCA 882
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
CG067437
LOCUS
DEFINITION
CG067437 Zea mays genomic clone ZMBR040592004,
genomic survey sequence.
ACCESSION
CG067437
VERSION
CG067437.1 GI:33939617
KEYWORDS
Zea mays
SOURCE
Zea mays
ORGANISM
Zea mays

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

# REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

707	CTTTAGCAAAATCCACCGGTGAACAAG---GGGGAGATAGTGTGTGACCGCATATCTCTA	763
651	TTCTTTTGCAAAATGCAACGTTGTGTGACGGCCAGTCTGACTCGGCGTTGGCCGATACCAATC	710
764	AAITTTTGATCTTGGACGGCACATGGATGTCCTCAGATTAAA	804
711	GCTTCGACCGTTAGATACACAGTGGACACTCTCGAATTAA	751

RESULT 15	LOCUS	582 bp	DNA	linear	GSS 06-NOV-2002
B212667	id37h06.b1	WGS -ZmaysF	JM107	adapted methyl filtered	Zea mays
DEFINITION	genomic clone id37h06.5	genomic survey sequence.			

SOURCE	ORGANISM	Zea mays	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 583)		
AUTHORS	Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N., Karzenburger, F., King, L., Miller, S., Muller, S., Nascimento, L., Zucaverni, T., McCombie, W.R. and Martenssen, R.A.		
TITLE	Genomic shotgun sequences from Zea mays (methyl-filtered)		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: W. Richard McCombie		

Query Match	8.3%	Score 207.8	DB 8	Length 582
Best Local Similarity	73.9%	Pred. No. 9.4e-38		
Matches 357; Conservative	0	Mismatches 117; Indels 9	Gaps 7	
242	GGTATAACAATCGCATGCTAAATAACACAAAGGAAATCTTACTAGAGCAGTAAATTTGGC	301		
51	GGTTTAGAAGCTTAACCAACAACACCCCTTAGGTATCTCTATAGGAGTAGTATTTGGC	110		
302	TAAACAATATGTGAGCATTTTATGTATATAGGGAATAGGAGCATGCAATATCTTGTTGTTCT	361		
111	CAAAACAATAATGAACA-TTTTAATGCAATGGGTGATACAGGAGTAGAATGTTTTCTTCT	169		
362	TTACGGGTTTGTATGCTCTCAAAAGTGTGCCCCCTGGGGCAGTTTGCAACACTCAAAATC	421		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 9, 2005, 05:52:07 ; Search time 1329 Seconds  
(without alignments)  
11367.636 Million cell updates/sec

Title: US-10-089-612-2  
Perfect score: 2493  
Sequence: 1 aagctttccggatgatgaag.....aaatagcttcacagtctaga 2493

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues  
Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCN\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCNUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73.4	2.9	602	18	US-10-425-115-24619 Sequence 24619, A
2	72.8	2.9	438	18	US-10-425-115-69828 Sequence 69828, A
3	72	2.9	495	18	US-10-425-115-93280 Sequence 93280, A
4	71.6	2.9	812	18	US-10-425-115-1000 Sequence 1000, A
5	62.4	2.5	2096	17	US-10-425-114-1206 Sequence 1206, A
6	62.4	2.5	2096	18	US-10-425-115-59907 Sequence 59907, A
7	62	2.5	8056	18	US-10-425-114-1206 Sequence 386, App
8	59	2.4	493	18	US-10-425-115-45869 Sequence 45869, A
9	58.8	2.4	8056	18	US-10-425-115-45869 Sequence 240, App
10	58.4	2.3	308	18	US-10-425-115-81221 Sequence 81221, A
11	56.4	2.3	1137	17	US-10-425-114-3317 Sequence 3317, A

C	12	56.4	2.3	1489	18	US-10-425-115-134984	Sequence 134984, A
	13	55.4	2.2	398	18	US-10-425-115-85377	Sequence 85377, A
	14	53.8	2.2	367378	16	US-10-312-841-1	Sequence 1, Appl
C	15	53.2	2.1	6863	17	US-10-221-714A-319	Sequence 319, App
C	16	52	2.1	12669	17	US-10-221-613-412	Sequence 412, App
	17	51.8	2.1	1157	18	US-10-425-115-161110	Sequence 161110, A
	18	51.8	2.1	5252	18	US-10-311-455-1108	Sequence 1108, App
C	19	51.8	2.1	9810	15	US-10-311-455-399	Sequence 399, App
	20	51.4	2.1	12142	15	US-10-311-455-1646	Sequence 1646, App
	21	51	2.0	15548	15	US-10-311-455-2128	Sequence 2128, App
	22	50.8	2.0	3252	13	US-10-027-632-113786	Sequence 113786, A
	23	50.8	2.0	3252	13	US-10-027-632-113787	Sequence 113787, A
	24	50.8	2.0	3252	13	US-10-027-632-113788	Sequence 113788, A
	25	50.8	2.0	3252	17	US-10-027-632-113786	Sequence 113786, A
	26	50.8	2.0	3252	17	US-10-027-632-113787	Sequence 113787, A
	27	50.8	2.0	3252	17	US-10-027-632-113788	Sequence 113788, A
C	28	50.6	2.0	7924	15	US-10-311-455-2112	Sequence 2112, App
C	29	50.6	2.0	7924	17	US-10-257-166-152	Sequence 152, App
C	30	50.4	2.0	13377	15	US-10-311-455-1436	Sequence 1436, App
C	31	50.4	2.0	13377	17	US-10-221-714A-198	Sequence 198, App
C	32	50.2	2.0	5488	15	US-10-311-455-1429	Sequence 1429, App
C	33	50.2	2.0	337344	18	US-10-388-838-58	Sequence 58, Appl
C	34	50	2.0	34688	18	US-10-433-793-90	Sequence 90, Appl
	35	49.8	2.0	7442	17	US-10-221-714A-409	Sequence 409, App
	36	49.6	2.0	113515	15	US-10-311-455-2148	Sequence 2148, App
	37	49.4	2.0	843	13	US-10-027-632-7034	Sequence 7034, App
	38	49.4	2.0	843	17	US-10-027-632-7034	Sequence 7034, App
C	39	49.4	2.0	367378	16	US-10-312-841-2	Sequence 2, Appl
	40	49.2	2.0	671	14	US-10-184-644-346	Sequence 346, App
	41	49.2	2.0	671	14	US-10-184-634-346	Sequence 346, App
C	42	48.8	2.0	2000	9	US-09-938-842A-4957	Sequence 4957, App
C	43	48.8	2.0	2000	11	US-09-938-842A-4957	Sequence 4957, App
C	44	48.8	2.0	4045	17	US-10-221-714A-466	Sequence 466, App
C	45	48.8	2.0	5238	15	US-10-311-455-731	Sequence 731, App

ALIGNMENTS

RESULT 1

US-10-425-115-24619  
; Sequence 24619, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 24619  
; LENGTH: 602  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_12245C.1  
US-10-425-115-24619

Query Match	2.9%	Score	73.4	DB	18	Length	602
Best Local Similarity	83.8%	Pred. No.	4.3e-06				
Matches	83	Conservative	0	Mismatches	16	Indels	0
Gaps	0						
Oy	1	AAGCTTTTCGGTGTGATGAAGCAGCTGTATCTACTTAAACAGCATCTGTAACAAATAGTTA	60				
Db	204	AAGCTTTTCGGTGTGATGAAGCAGCTGTATCTACTTAAACAGCATCTGTAACAAATAGTTA	263				
Oy	61	GCTGTGTTTTGAGGACCTTCGGAAGATGAAGGCCCA	99				
Db	264	GCTGTGTTTTGAGGACCTTCGGAAGATGAAGGCCCA	302				



Oy 61 GCTGTGTTTTCAGGACCTTCGGAAGATGAAGGCCGCCCAACATCCCATCAAGTC 120  
Db 1709 ACCGTGTTTT-AAGACCTTCAGAGAGAGAGGCCGCCCAAGACCACTGAACCTGATGAC 1651  
Oy 121 CCATGACCTTCGCAAAAAGCAATTTTA 148  
Db 1650 ACAACATTAATGAAGACATCTCATGCTA 1623

## RESULT 6

US-10-425-115-59907/c  
; Sequence 59907, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 59907  
; LENGTH: 2096  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_154633C.1  
US-10-425-115-59907

Query Match 2.5%; Score 62.4; DB 18; Length 2096;  
Best Local Similarity 68.2%; Pred. No. 0.0025;  
Matches 101; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

Oy 1 AGCGTTTCGGTGATGAGACCTGTAATCTTAACAGCATGCTGAACCAATAGTTA 60  
Db 1769 AAGTTTTTCTAGTACAGAGCTCTGTGTAATCTTAATACATCTGAATAATGTTA 1710  
Oy 61 GCTGTGTTTTCAGGACCTTCGGAAGATGAAGGCCGCCCAACATCCCATCAAGTC 120  
Db 1709 ACCGTGTTTT-AAGACCTTCAGAGAGAGAGGCCGCCCAAGACCACTGAACCTGATGAC 1651  
Oy 121 CCATGACCTTCGCAAAAAGCAATTTTA 148  
Db 1650 ACAACATTAATGAAGACATCTCATGCTA 1623

## RESULT 7

US-10-473-126-386/c  
; Sequence 386, Application US/10473126  
; Publication No. US20040214973A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell  
; FILE OF INVENTION: Proliferative disorders  
; FILE REFERENCE:  
; CURRENT APPLICATION NUMBER: US/10/473,126  
; CURRENT FILING DATE: 2003-09-26  
; NUMBER OF SEQ ID NOS: 1258  
; SEQ ID NO 386  
; LENGTH: 8056  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-473-126-386

Query Match 2.5%; Score 62; DB 18; Length 8056;  
Best Local Similarity 45.5%; Pred. No. 0.0062;  
Matches 221; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Oy 1852 AATTGAAGTGAATTTTAAATCAATCAATGATCCCAATCTCTCCAAATATACCAA 1911  
Db 2361 AATTATATTTAAATAATTTAAATTAATTTTAAATACAAATTTAAATTAATTTTAA 2302  
Oy 1912 ATATGAATATTTAGATGATATGTTGGTGGAGTTTGGGCTCCGCTTTGGTTAGTATCT 1971  
Db 2301 AAATAAATAATTTAAATAATTTTAAATTTTAAATTTTAAATAATTTAAATAAAT 2242  
Oy 1972 TTGTATAAAAATTAATTTCTCTCTCTTTGTCTCACTTCCCAATTTGACTTAAAAATTTTATGTA 2031  
Db 2241 TTTTATTAATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2182  
Oy 2032 GCATGCCCACTTTTCTAGTGTGCCACTTATAGCACAAACAACTATATCCATTTCT 2091  
Db 2181 ATAAATTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATAAATTTATTA 2122  
Oy 2092 AATAGTCTTGAAATCCCACTTCTATTTTGTAGCCATTTCTCAAAATTTGGCACAACAACTAG 2151  
Db 2121 AATTATATTAATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATAAATTT 2062  
Oy 2152 GAAATTTAATACATTTCTGCCATACATATTTCTAGTGCAAATGTTAACTAGATGCTCA 2211  
Db 2061 TTTTAAATAATTAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAT 2002  
Oy 2212 ATATTAGCAACCTCTTTTCTAGATTCATTAATTTGCTACATTTGCTATGCTATCTTTTATGA 2271  
Db 2001 AATTTTAAATAAATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTATTT 1942  
Oy 2272 AGTTTCATCAATGCTCTCATTTAGCATCTCTTATTTAGCACTTGAATTAACCCGCTT 2331  
Db 1941 ATATTATTAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATAAATA 1882  
Oy 2332 AAAATA 2337  
Db 1881 AAAATA 1876

## RESULT 8

US-10-425-115-45669  
; Sequence 45669, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(5322)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 45669  
; LENGTH: 493  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_141656C.1  
US-10-425-115-45669

Query Match 2.4%; Score 59; DB 18; Length 493;  
Best Local Similarity 78.0%; Pred. No. 0.0068;  
Matches 71; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Oy 23 CCTGTATATTTAAGCATGCTGGAACAAATAGTTAGCTGTGTTTTGAGGACCTTGG 82  
Db 370 CCTGTATAGTCTATATCATCTGGAACAAATAGTTAGTACATTTTGGACCTTGG 429  
Oy 83 GAAGATGAAGGCCGCCCAACATCCCATGCA 113  
Db 430 GAAGAGGAGGCCCTCAACACATTAATCTGGA 460

```
RESULT 9
US-10-473-126-240/c
; Sequence 240, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 240
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically created genomic DNA (Homo sapiens)
US-10-473-126-240

Query Match      2.4%; Score 58.8; DB 18; Length 8056;
Best Local Similarity 45.1%; Pred. No. 0.032;
Matches 219; Conservative 0; Mismatches 267; Indels 0; Gaps 0;

Qy 1852 AATTTAAGTTCGAATTTTAAATTCAAATCAATGACCTCCAAATCTCTCCAAATACCAAA 1911
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2361 AATATATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2302

Qy 1912 ATATAGAAATATTAGATGAATATGTGTGGAGTGTGGGCTCCGCTTTGGTTAGTAGTGT 1971
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2301 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2242

Qy 1972 TTGTATAAATAAATAATTTCTCTCTCTTTGTGCACTTCCAAATATGACTTAATTTTATGTA 2031
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2241 TTTATATAAATAAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2182

Qy 2032 GCAATGCCAACTTTTTTTAGTAGTGCCACTTATAGCACAAAACATATATCCATTTCT 2091
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2181 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2122

Qy 2092 AATAGTCTTGAATCCACATCTATTTTAGCCATCTTCAAAATGGCCACAAACAG 2151
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2121 AATATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2062

Qy 2152 GAAATTTAATACATCTTCCCAATACATATTTCTAGTGCATATGTTAACTAGATGCTCA 2211
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2061 TTTTAAAAATAAAAAATAATTTAAAAAATAATTTTATTTTAAAAATATATTTAAAAATTTTAT 2002

Qy 2212 ATATTAGCAACTCTTTTGTAGATTCATTAATATGCTACATTCATATCTTTTATAGA 2271
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2001 AATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1942

Qy 2272 AGTTCAATCAATAGCCTCATTTAGCATCTTCATTTTAGGAATCTTGATTAACCGCCTT 2331
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1941 ATATTATTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1882

Qy 2332 AAAATA 2337
      |||||
Db 1881 AAAATA 1876
      |||||

RESULT 10
US-10-425-115-81221/c
; Sequence 81221, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3317
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700259576_FU1
US-10-425-114-3317

Query Match      2.3%; Score 56.4; DB 17; Length 1137;
Best Local Similarity 80.5%; Pred. No. 0.04;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 23 CCTGTAAATCTTAACAGCATGCTGAAAAACAAATAGTTAGCTGTGTTTTGAGGACCTTCG 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 778 CCTGTAAATCTCACTATTATGCTAGAACAAATGTTAGTCATGTTTTGAGGACCTTAG 719

Qy 83 GAAGATGAAGGCCCCCAACACA 104
      |||||
Db 718 AAAGAGGAAGGCCCCCAACAAA 697
      |||||

RESULT 12
US-10-425-115-134984/c
; Sequence 134984, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3317
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700259576_FU1
US-10-425-114-3317

Query Match      2.3%; Score 56.4; DB 17; Length 1137;
Best Local Similarity 80.5%; Pred. No. 0.04;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 23 CCTGTAAATCTTAACAGCATGCTGAAAAACAAATAGTTAGCTGTGTTTTGAGGACCTTCG 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 778 CCTGTAAATCTCACTATTATGCTAGAACAAATGTTAGTCATGTTTTGAGGACCTTAG 719

Qy 83 GAAGATGAAGGCCCCCAACACA 104
      |||||
Db 718 AAAGAGGAAGGCCCCCAACAAA 697
      |||||

RESULT 11
US-10-425-114-3317/c
; Sequence 3317, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3317
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700259576_FU1
US-10-425-114-3317

Query Match      2.3%; Score 58.4; DB 18; Length 308;
Best Local Similarity 85.5%; Pred. No. 0.0072;
Matches 65; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 27 TAATACTTAACAGCATGCTGAAAAACAAATAGTTAGCTGTGTTTTGAGGACCTTCGGAAG 86
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 TAATACTTCAATAATATCTGTAACAAACAGATGTTAGCGGTATTTAAGGACCTTCGAAAA 249
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 87 ATGAAGGCCCCCAACA 102
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 ATGAAGGCCCCCAACA 233
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
US-10-425-114-3317/c
; Sequence 3317, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3317
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700259576_FU1
US-10-425-114-3317

Query Match      2.3%; Score 56.4; DB 17; Length 1137;
Best Local Similarity 80.5%; Pred. No. 0.04;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 23 CCTGTAAATCTTAACAGCATGCTGAAAAACAAATAGTTAGCTGTGTTTTGAGGACCTTCG 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 778 CCTGTAAATCTCACTATTATGCTAGAACAAATGTTAGTCATGTTTTGAGGACCTTAG 719

Qy 83 GAAGATGAAGGCCCCCAACACA 104
      |||||
Db 718 AAAGAGGAAGGCCCCCAACAAA 697
      |||||

RESULT 12
US-10-425-115-134984/c
; Sequence 134984, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3317
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700259576_FU1
US-10-425-114-3317

Query Match      2.3%; Score 56.4; DB 17; Length 1137;
Best Local Similarity 80.5%; Pred. No. 0.04;
Matches 66; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 23 CCTGTAAATCTTAACAGCATGCTGAAAAACAAATAGTTAGCTGTGTTTTGAGGACCTTCG 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 778 CCTGTAAATCTCACTATTATGCTAGAACAAATGTTAGTCATGTTTTGAGGACCTTAG 719

Qy 83 GAAGATGAAGGCCCCCAACACA 104
      |||||
Db 718 AAAGAGGAAGGCCCCCAACAAA 697
      |||||

RESULT 11
US-10-425-114-3317/c
; Sequence 3317, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3317
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700259576_FU1
US-10-425-114-3317

Query Match      2.3%; Score 58.4; DB 18; Length 308;
Best Local Similarity 85.5%; Pred. No. 0.0072;
Matches 65; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 27 TAATACTTAACAGCATGCTGAAAAACAAATAGTTAGCTGTGTTTTGAGGACCTTCGGAAG 86
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 TAATACTTCAATAATATCTGTAACAAACAGATGTTAGCGGTATTTAAGGACCTTCGAAAA 249
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 87 ATGAAGGCCCCCAACA 102
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 ATGAAGGCCCCCAACA 233
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 9, 2005, 02:16:11 : Search time 432 Seconds  
(without alignments)  
9442.680 Million cell updates/sec

Title: US-10-089-612-2

Perfect score: 2493

Sequence: 1 agctttccggtgatgaag.....aaatagcttcacagctctaga 2493

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	74	3.0	7218	1	US-08-232-463-14
C 2	56	2.2	1141	4	US-09-806-708B-22
C 3	55.2	2.2	30820	4	US-09-949-016-17145
C 4	51.2	2.1	1664976	4	US-08-916-421B-1
C 5	51.2	2.1	1664976	4	US-09-692-570-1
C 6	48.2	1.9	601	4	US-09-949-016-30530
C 7	48.2	1.9	601	4	US-09-949-016-30531
C 8	48.2	1.9	601	4	US-09-949-016-37149
C 9	48.2	1.9	601	4	US-09-949-016-37150
C 10	48.2	1.9	601	4	US-09-949-016-37163
C 11	48.2	1.9	601	4	US-09-949-016-37164
C 12	48.2	1.9	601	4	US-09-949-016-145867
C 13	48.2	1.9	601	4	US-09-949-016-145868
C 14	48.2	1.9	601	4	US-09-949-016-146135
C 15	48.2	1.9	601	4	US-09-949-016-146136
C 16	48.2	1.9	601	4	US-09-949-016-146403
C 17	48.2	1.9	601	4	US-09-949-016-146404
C 18	48.2	1.9	205044	4	US-09-949-016-15851
C 19	48.2	1.9	205044	4	US-09-949-016-15852
C 20	48.2	1.9	205044	4	US-09-949-016-15853
C 21	48.2	1.9	223471	4	US-09-949-016-12387
C 22	48.2	1.9	223471	4	US-09-949-016-12724
C 23	48.2	1.9	223471	4	US-09-949-016-12725
C 24	48	1.9	251682	4	US-09-949-016-17296
C 25	48	1.9	251682	4	US-09-949-016-11973
C 26	47.6	1.9	19124	2	US-08-487-836B-13
C 27	46.6	1.9	2121	4	US-09-248-796A-4904

C 28	46.4	1.9	67755	4	US-09-949-016-13703	Sequence 13703, A
C 29	46.2	1.9	601	4	US-09-949-016-54020	Sequence 54020, A
C 30	46.2	1.9	601	4	US-09-949-016-108654	Sequence 108654, A
C 31	46.2	1.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
C 32	46	1.8	660	1	US-07-991-867B-32	Sequence 32, Appl
C 33	46	1.8	660	1	US-08-107-755A-32	Sequence 32, Appl
C 34	46	1.8	660	2	US-08-544-332-32	Sequence 32, Appl
C 35	46	1.8	660	3	US-09-370-861A-32	Sequence 32, Appl
C 36	46	1.8	1511	1	US-07-991-867B-8	Sequence 8, Appl
C 37	46	1.8	1511	1	US-08-107-755A-8	Sequence 8, Appl
C 38	46	1.8	1511	2	US-08-544-332-8	Sequence 8, Appl
C 39	46	1.8	1511	3	US-09-370-861A-8	Sequence 8, Appl
C 40	46	1.8	4810	3	US-08-852-629-11	Sequence 11, Appl
C 41	46	1.8	4838	3	US-08-852-629-15	Sequence 15, Appl
C 42	45.6	1.8	3848	2	US-08-808-931-14	Sequence 14, Appl
C 43	45.6	1.8	3848	3	US-08-808-323-14	Sequence 14, Appl
C 44	45.6	1.8	3848	3	US-09-050-603A-14	Sequence 14, Appl
C 45	45.6	1.8	3848	3	US-09-102-420B-14	Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-08-232-463-14/c  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
CLONE: pTZgpt-F18  
US-08-232-463-14

Query Match 3.0%; Score 74; DB 1; Length 7218;





LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664855)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-4218-1

Query Match 2.1%; Score 51.2; DB 4; Length 1664976;  
Best Local Similarity 44.9%; Pred. No. 0.076;  
Matches 240; Conservative 0; Mismatches 288; Indels 6; Gaps 1;

Qy 1958 TTGGTTAGTATGTTTGTATATAAATAATTTCTCTCTTTCTGCTTCCATTTGACT 2017  
Db 248153 TTTCTCTATTTTAAATTTTCCAACTTTTATTAATTTCTATATAAACCCTTTATCCATT 248094

Qy 2018 TAAATTTTATGTCAGCAATGCCAACTTTTTTTAGTAGTGGCACTTATAGCACAAAAC 2077  
Db 248093 AAATCACCAGTTGCCATATATAATTTTAACTTTTAACTTTTATAGAAATAATTT 248034

Qy 2078 TATATCATTTTCTAATAGTCTTGAATCCACATCTCTTTTAGCCATCTTCAAAAT 2137  
Db 248033 TAAAGTATTTATAGGTATCCATTCACCTATATAAATTTCTCTGTTTCAAGTATAT 247974

Qy 2138 TGGCACAAAAC TAGGAAAATTTAATACATTTCTGCCATACACATATTTCTAGTGCAAATGTT 2197  
Db 247973 TGCTTAAATTTAATAATTTTATATAGTTTAAATATAAATAATTTGT-----ATAATTT 247920

Qy 2198 AACTAGATTGCTCAATATTAGCAAACTCTTTTGTAGATTCATTAATTTGCTACATTTG 2257  
Db 247919 AAATATATAGGTGAAAATATGAATTTAATATTAACCTCTCATTAATTTTGGCATTAAT 247860

Qy 2258 CATACTTTTTTGAAGTTCATCAATATGCTCATATTAGCATCTTCAATTTTGGAACTTG 2317  
Db 247859 GGGCTTTTAAATTTTAAATTTTGGAAATTTGGCTGATAAAAAAACTTCTATTTCCAAAAAG 247800

Qy 2318 ATTAAACCCCTTAAATAGACCAAGTGACCGATCCATTTAAAGGTGATTTCTTAATTT 2377  
Db 247799 AAGCCAAACCATATCAAGTTTAGACATGGAAGCTCATCAAGTAAAGGGTTATGGAGT 247740

Qy 2378 CTACTTCTATCTTTGGTGGCTTATTTATATATATGTTGGGTGTTCAATGATGTTCC 2437  
Db 247739 CATGATGTTTATCATCACCATTATGTAACCTAAAGATGTTCTATGTCATGATTTATGGCAT 247680

Qy 2438 TACACCACTACCCACACGTTGGACATATATATGAAAAATAGCTTCCACAGTCTA 2491  
Db 247679 GACGACATTTATGTCGAAGAGGAAAGAGGGGTGATAAAGATGATAGTTTA 247626

RESULT 5  
US-09-692-570-1/c  
Sequence 1, Application US/09692570  
Patent No. 6797466  
GENERAL INFORMATION:  
APPLICANT: Rult et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
FILE REFERENCE: PB275C1  
CURRENT APPLICATION NUMBER: US/09/692,570  
CURRENT FILING DATE: 2003-01-14  
PRIOR APPLICATION NUMBER: US 60/024,428  
PRIOR FILING DATE: 1996-08-22  
PRIOR APPLICATION NUMBER: US 08/916,421  
PRIOR FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
FEATURE:

```

; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature

```

```

; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

```

Query Match 2.1t; Score 51.2; DB 4; Length 1664976;  
Best Local Similarity 44.9t; Pred. NO. 0.076;



Matches	240;	Conservative	0;	Mismatches	288;	Indels	6;	Gaps	1;
Qy	1958	TTTGTTAGTATGTTGTGATATAAAAAATAATTTCTCTCTCTTTTGTGCATCTTCCAAATATGTACT	2017						
Db	248153	TTTTCTCAITTTTAAATTTTCAACCTTTTTTTATTTAAATCTATAAAACCTCTTATCCATT	248094						
Qy	2018	TAAATTTTTATGTAGCAATGCCAACTTTTTTTAGTAGTGTGCCACTTTATAGCACAAAAAC	2077						
Db	248093	AAATCAACCAATGCCAATAATAAATTTTTTAATCTTTAACTATTTATAGAAATATTT	248034						
Qy	2078	TATATCCATTTTCTAATPAGTCTTGAATCCAATCTTAATTTTTTAGGCCATCTTCCAAAT	2137						
Db	248033	TAAAGTAITTTATAGGTATCCATTCAACTATATAAAAAATTTCTACTGATTCAGGTATAT	247974						
Qy	2138	TGCGACAAATCTAGGAAATTTAATACATTTCTTGCCATAACATATTTCTAGTGCAAATGTT	2197						
Db	247973	TGCTTAAGATTTTAAATTTTTTATATAGTTTAAATATATAAANATTTGT-----ATAATTT	247920						
Qy	2198	AACAGATGCTCAATATATAGCAAACTTCTTTTGTAAAGATTCATTAAATATGCTACATGT	2257						
Db	247919	AAATATATAGTGGAAATATGAATTTAATATTAACCTACTACTCAATATTTTGCCTAAAT	247860						
Qy	2258	CATACTTTTTTAGAAAGTTCAATCAATATGCTCATTTAGCATACTTCATTTTAGGAAGCTTG	2317						
Db	247859	GGGCTTTTAAITTTTAAITTTTGGAAATTTTGGCTATATAAAAACCTCTTATTTCCAAAAAG	247800						
Qy	2318	ATTAAACCGCGCTTAAATATAGACGCCAGTCAACGGATCCAATTTAAAGGTGAITTTCTTAATTT	2377						
Db	247799	AAGCCAAACCATATCAAGTTTAGACATATGGAAGCTCATCAAGTAAAGAGGTTATGGAAGT	247740						
Qy	2378	CTTACTTCTCATTTTGGTGGCTTANGTTTTATATATGTGGGTGGTGGTGAATGATGTTC	2437						
Db	247739	CATGATGTTTATCATCAACCATATATGTAACATAAAGATGTCTATGTCCCAATGATTATGGGAT	247680						
Qy	2438	TACACCATCTACACCAACGTTGGACATATATATGGAATATAGCTTCCAGTCTCTA	2491						
Db	247679	GAGACATTTATGATGAAGAGGAAAAAAGAGGGTGATAAAGATGATAGCTTTA	247626						

## RESULT 6

```

US-09-949-016-30530/c
; Sequence 30530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fasmseq for Windows Version 4.0
; SEQ ID NO 30530
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30530

```

## RESULT 8

US-09-949-016-37149/c  
; Sequence 37149, Application US/09949016  
; Patent NO. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37149
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37149

Query Match
  1.9%; Score 48.2; DB 4; Length 601;
Best Local Similarity 51.1%; Pred. No. 0.0047;
Matches 113; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 429 ATACATAAAGAAACATGGGCACCAAAATAGAAACAATCTCAAAATTTAGACCAAAACCACTTTAAAT 548
   |||||
Db 380 ATATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 321

Qy 489 AATGGCTCTAATATTTAGTACATTTTGTAGATTTTATTTAGACCAAAACCACTTTAAAT 548
   |||||
Db 320 ATTTAAATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATAT 261

Qy 549 TTGGTTTAAATGAGTTAGATATATATATATATATATATATATATATATATATATATATATATAT 608
   |||||
Db 260 ATATTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 201

Qy 609 TTTACTTAACATACTCTAGGGTTTAAAGTAAATTTT 649
   |||||
Db 200 AATATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATAT 160

RESULT 9
US-09-949-016-37150/c
; Sequence 37150, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37150
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37150

Query Match
  1.9%; Score 48.2; DB 4; Length 601;
Best Local Similarity 51.1%; Pred. No. 0.0047;
Matches 113; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 429 ATACATAAAGAAACATGGGCACCAAAATAGAAACAATCTCAAAATTTAGACCAAAACCACTTTAAAT 548
   |||||
Db 380 ATATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 321

Qy 489 AATGGCTCTAATATTTAGTACATTTTGTAGATTTTATTTAGACCAAAACCACTTTAAAT 548
   |||||
Db 320 ATTTAAATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATAT 261

Qy 549 TTGGTTTAAATGAGTTAGATATATATATATATATATATATATATATATATATATATATATATAT 608
   |||||
Db 260 ATATTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 201

Qy 609 TTTACTTAACATACTCTAGGGTTTAAAGTAAATTTT 649
   |||||
Db 200 AATATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATAT 160

RESULT 10
US-09-949-016-37163/c
; Sequence 37163, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37163
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37163

Query Match
  1.9%; Score 48.2; DB 4; Length 601;
Best Local Similarity 51.1%; Pred. No. 0.0047;
Matches 113; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 429 ATACATAAAGAAACATGGGCACCAAAATAGAAACAATCTCAAAATTTAGACCAAAACCACTTTAAAT 548
   |||||
Db 380 ATATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 321

Qy 489 AATGGCTCTAATATTTAGTACATTTTGTAGATTTTATTTAGACCAAAACCACTTTAAAT 548
   |||||
Db 320 ATTTAAATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATAT 261

Qy 549 TTGGTTTAAATGAGTTAGATATATATATATATATATATATATATATATATATATATATATATAT 608
   |||||
Db 260 ATATTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 201

Qy 609 TTTACTTAACATACTCTAGGGTTTAAAGTAAATTTT 649
   |||||
Db 200 AATATATAAATAATATATATATATATATATATATATATATATATATATATATATATATATAT 160

RESULT 11
US-09-949-016-37164/c
; Sequence 37164, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37164
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-37164
```



Db 380 ATATATAAAATATATTTTATATATATTTTAAATATAATATAATATAATATATATAT 321  
Qy 489 AATGGCTCTAATATATAGACATTTAGAAATTTATTTAGACCAAAACCATTTAAAT 548  
Db 320 ATTTTAAATATAATATAATATAATATATATATATTTTAAATATAATATAATATAAT 261  
Qy 549 TTGGTTTAAATAGTAGATATATATATTTTATTCAGTTTATAGTTATTTGGGACATTTA 608  
Db 260 ATATTTATATATTTTAAATATAATATAATATAATATAATATAATATAATATAATATA 201  
Qy 609 TTACTTAACTATACCTCTAGGGTTTAAAGTAAATTTT 649  
Db 200 AATATAAAATATATTTTATATATATTTTAAATATAATATAATATAAT 160

## RESULT 15

US-09-949-016-146136/c  
; Sequence 146136, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 146136  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-146136

Query Match 1.9%; Score 48.2; DB 4; Length 601;  
Best Local Similarity 51.1%; Pred. No. 0, 0047;  
Matches 113; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
Qy 429 ATACATAAGAAACATGGGCACAAAATAGAAACATCTCAAAATTTAGAAAAGTTCA 488  
Db 470 ATATATAAAATATATTTTATATATATTTTAAATATAATATAATATAATATAATAT 411  
Qy 489 AATGGCTCTAATATTTAGACATTTTAGAAATTTTATTTAGACCAAAACCATTTAAAT 548  
Db 410 ATTTTAAATATAATATAATATAATATATTTTATATATATTTTAAATATAATATAAT 351  
Qy 549 TTGGTTTAAATAGTAGATATATATATTTTATTCAGTTTATAGTTATTTGGGACATTTA 608  
Db 350 ATATTTATATATTTTAAATATAATATAATATAATATAATATAATATAATATAATATA 291  
Qy 609 TTACTTAACTATACCTCTAGGGTTTAAAGTAAATTTT 649  
Db 290 AATATAAAATATATTTTATATATATTTTAAATATAATATAATATAAT 250

Search completed: April 9, 2005, 07:59:45  
Job time : 438 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 20:49:49 / Search time 1154 Seconds  
(without alignments)  
12788.488 Million cell updates/sec

Title: US-10-089-612-2

Perfect score: 2493

Sequence: 1 agctttccggtgatgaag.....aatagcttcacagctctaga 2493

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: N\_Geneseq\_16Dec04.\*
- 2: Geneseqm1980s.\*
- 3: Geneseqm1990s.\*
- 4: Geneseqm2000s.\*
- 5: Geneseqm2001as.\*
- 6: Geneseqm2001bs.\*
- 7: Geneseqm2002as.\*
- 8: Geneseqm2002bs.\*
- 9: Geneseqm2003as.\*
- 10: Geneseqm2003bs.\*
- 11: Geneseqm2003ds.\*
- 12: Geneseqm2004as.\*
- 13: Geneseqm2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2493	100.0	2493	4 AAF80481	AAF80481 Nucleotid
2	499	20.0	499	4 AAF80484	AAF80484 Nucleotid
3	341	13.7	1708	4 AAF80482	AAF80482 Nucleotid
4	339	13.6	507	4 AAF80485	AAF80485 Nucleotid
5	236.8	9.5	1232	4 AAF80483	AAF80483 Nucleotid
6	216.2	8.7	265	4 AAF80486	AAF80486 Consensus
7	163.6	6.6	531	4 AAF80480	AAF80480 Nucleotid
8	62	2.5	8056	8 ABZ10246	ABZ10246 Haematopo
9	60.4	2.4	10439	6 AAD47817	AAD47817 Alternati
10	60.4	2.4	11779	6 AAD42891	AAD42891 Maize sug
11	58.8	2.4	8056	8 ABZ10100	ABZ10100 Haematopo
12	55.6	2.2	2000	8 ADA71938	ADA71938 Rice gene
13	53.2	2.1	6863	4 AAS46617	AAS46617 Tumour eu
14	52	2.1	12669	6 ABK31523	ABK31523 Signal tr
15	52	2.1	12669	6 ABL70490	ABL70490 Chemical
16	52	2.1	12669	6 AAS61449	AAS61449 Human gen
17	51.8	2.1	5252	6 ABL33135	ABL33135 Human inm
18	51.8	2.1	9810	6 ABL32426	ABL32426 Human inm
19	51.4	2.1	12142	6 ABL33673	ABL33673 Human inm
20	51.2	2.1	347	4 AAI87539	AAI87539 Human pol

c 21	51.2	2.1	110000	2 AAV21209_02	Continuation (3 of
c 22	51	2.0	15548	6 ABL34155	ABL34155 Human inm
c 23	50.6	2.0	7924	6 ABK40070	ABK40070 Human che
c 24	50.6	2.0	7924	6 ABL34139	ABL34139 Human inm
c 25	50.4	2.0	13377	4 AAS46476	AAS46476 Tumour eu
c 26	50.4	2.0	13377	6 ABL33463	ABL33463 Human inm
c 27	50.2	2.0	5488	6 ABL33456	ABL33456 Human inm
c 28	50.2	2.0	33734	13 ABD32715	ABD32715 Human can
c 29	50	2.0	4001	13 ADS89759	ADS89759 Oligonuc1
c 30	50	2.0	8900	13 ADS89686	ADS89686 Oligonuc1
c 31	50	2.0	34688	6 ABQ67060	ABQ67060 Human ang
c 32	49.8	2.0	7442	4 AAS46686	AAS46686 Tumour eu
c 33	48.6	2.0	113515	6 ABL34175	ABL34175 Human inm
c 34	48.8	2.0	2000	6 ABZ17152	ABZ17152 Arabidops
c 35	48.8	2.0	4045	4 AAS46740	AAS46740 Tumour eu
c 36	48.8	2.0	5238	6 ABL32758	ABL32758 Human inm
c 37	48.8	2.0	110000	8 AAS53224_2	Continuation (3 of
c 38	48.6	1.9	5666	6 AAS61186	AAS61186 Human gen
c 39	48.6	1.9	5666	6 ABL49329	ABL49329 Human pol
c 40	48.4	1.9	4001	13 ADS89485	ADS89485 Oligonuc1
c 41	48.4	1.9	6167	6 ABK28370	ABK28370 DNA trans
c 42	48.4	1.9	6534	6 ABL32470	ABL32470 Human inm
c 43	48.4	1.9	8900	13 ADS89412	ADS89412 Oligonuc1
c 44	48.4	1.9	83391	6 ABQ67093	ABQ67093 Human ang
c 45	48.2	1.9	50000	6 ABL55643	ABL55643 AmSPV gen

## ALIGNMENTS

RESULT 1  
AAF80481  
ID AAF80481 standard; DNA; 2493 BP.  
XX  
AC AAF80481;  
DT 25-JUN-2001 (first entry)  
XX  
DE Nucleotide sequence of the promoter of the Esr2 gene of maize.  
XX  
KW Promoter; Esr2 gene; albumen; Angiosperm seed; embryo development;  
KW transgenic plant; seed starch; seed oil; food; ss.  
XX  
OS Zea mays.  
XX  
FN FR2799203-Al.  
XX  
PD 06-APR-2001.  
XX  
PF 01-OCT-1999; 99FR-00012305.  
XX  
PR 01-OCT-1999; 99FR-00012305.  
XX  
PA (BIOG-) BIOGEMMA GAS.  
XX  
PI Bonello JF, Rogowsky P, Perez P;  
XX  
DR WPI; 2001-275283/29.  
XX  
PT New plant promoters providing albumen-specific expression, useful for  
PT preparing transgenic plants with altered starch and oil contents.  
XX  
PS Claim 3; Page 41-42; 47pp; French.  
XX  
CC The present sequence represents a promoter of the Esr2 gene of maize. The  
CC promoter provides expression of attached sequences that is specific to  
CC the region of albumen surrounding the embryo in angiosperm seeds and  
CC active in the early stages of albumen development. The promoter provides  
CC specific expression of genes at an early stage of embryo development,  
CC resulting in seeds of increased size, enriched in starch and/or oil, or  
CC embryo-free seeds. The promoter is used to prepare expression cassettes  
CC for preparation of transgenic Angiosperm plants with improved agronomic  
CC or nutritional properties. Especially, these plants produce seeds with

CC	altered starch and oil contents and are useful as sources of human and	
CC	animal foods	
XX		
SQ	Sequence 2493 BP; 738 A; 430 C; 597 G; 728 T; 0 U; 0 Other:	
	Query Match 100.0%; Score 2493; DB 4; Length 2493;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2493; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AAGCTTTTCGGGTGATGAAGACACCTGTAATACAGACATGCTGAAACAAATAGTTA 60	
DB	1 AAGCTTTTCGGGTGATGAAGACACCTGTAATACAGACATGCTGAAACAAATAGTTA 60	
QY	61 GCTGTGTTTTGAGGACCTTGGGAAGATGAAGGCCCCCAACACATCCCATGATCAAGTC 120	
DB	61 GCTGTGTTTTGAGGACCTTGGGAAGATGAAGGCCCCCAACACATCCCATGATCAAGTC 120	
QY	121 CCATGACCTTGCAAAAGCAAAATTTATCAAAATTTCTATAAACACCTTGAAAACATT 180	
DB	121 CCATGACCTTGCAAAAGCAAAATTTATCAAAATTTCTATAAACACCTTGAAAACATT 180	
QY	181 TCTCTTTTGAAGGTGTAGAGCACTAGCAACTGCTCTACTAAAGGTTCCCAATTTCT 240	
DB	181 TCTCTTTTGAAGGTGTAGAGCACTAGCAACTGCTCTACTAAAGGTTCCCAATTTCT 240	
QY	241 GGGTAAACATCGCATGTTAAATACAAAGGAATCTTACTTAAGAGCAGTAATTTGG 300	
DB	241 GGGTAAACATCGCATGTTAAATACAAAGGAATCTTACTTAAGAGCAGTAATTTGG 300	
QY	301 CTAAACATAGTGAAGCATTTTAAATAGGATAGGAGCATGATCTGCTGTTTC 360	
DB	301 CTAAACATAGTGAAGCATTTTAAATAGGATAGGAGCATGATCTGCTGTTTC 360	
QY	361 TTTTCAGGGTTTGTATGCTCTCAAAAGTGTGCCCTCGGGCAGTGTGCAACACTCAAAAT 420	
DB	361 TTTTCAGGGTTTGTATGCTCTCAAAAGTGTGCCCTCGGGCAGTGTGCAACACTCAAAAT 420	
QY	421 CTACTGTTATACATTAAGGAACAATGGGCAAAATTAAGGAACAATTAAGGAACAATTAAGGA 480	
DB	421 CTACTGTTATACATTAAGGAACAATGGGCAAAATTAAGGAACAATTAAGGAACAATTAAGGA 480	
QY	481 AAGGTTCAAAATGGTCTTATTAATTTATGAGACATTTTAAAGTTTATTTAGACCAAAACC 540	
DB	481 AAGGTTCAAAATGGTCTTATTAATTTATGAGACATTTTAAAGTTTATTTAGACCAAAACC 540	
QY	541 ATTAAATTTGGTTTAAATGAGTTAGATATTAATTTATTTAGATTTATGATTTATTTGG 600	
DB	541 ATTAAATTTGGTTTAAATGAGTTAGATATTAATTTATTTAGATTTATGATTTATTTGG 600	
QY	601 GACATTTATTTACTTAACATTAACCTTCTAGGGTTTAAAGTTAAATTTTGGGTCCTTAGT 660	
DB	601 GACATTTATTTACTTAACATTAACCTTCTAGGGTTTAAAGTTAAATTTTGGGTCCTTAGT 660	
QY	661 TGGACCTAGCTCAGATTTGCTGTTGATTTTCCATAAAGTCCAGGTTCTTTTAGCAAAAT 720	
DB	661 TGGACCTAGCTCAGATTTGCTGTTGATTTTCCATAAAGTCCAGGTTCTTTTAGCAAAAT 720	
QY	721 CCAAGGTGAACAAAGGGGAGATAGGTGTTGACCGATATCTCTAAATTTGATCGTTGAC 780	
DB	721 CCAAGGTGAACAAAGGGGAGATAGGTGTTGACCGATATCTCTAAATTTGATCGTTGAC 780	
QY	781 GGCACATGATGCTCTCAGATTTAAATGGTGGATGTGCAAGCGACGCGACACATGAGGA 840	
DB	781 GGCACATGATGCTCTCAGATTTAAATGGTGGATGTGCAAGCGACGCGACACATGAGGA 840	
QY	841 ATGGCTCTACGACGGTGGCTACTAGAGCTGGCTAGTCAACCAATGGGGCTCGGTCA 900	
DB	841 ATGGCTCTACGACGGTGGCTACTAGAGCTGGCTAGTCAACCAATGGGGCTCGGTCA 900	
QY	901 AGGTCAAAATTTGTTGGCAAGCCACTGTGGCTCACGATCGAGTTCGATTCGACATATCAA 960	
DB	901 AGGTCAAAATTTGTTGGCAAGCCACTGTGGCTCACGATCGAGTTCGATTCGACATATCAA 960	

QY	961 GGTCCAGGGCTCAACAGAGGGGCAAGATCGATGTCGATGCTGCTCGATGGAAGGG 1020	
DB	961 GGTCCAGGGCTCAACAGAGGGGCAAGATCGATGTCGATGCTGCTCGATGGAAGGG 1020	
QY	1021 AAGCTTCGGGTGAGCAATTCAGATTTCTTATCATGTCACCGGTCAAGGAATGGGCCAT 1080	
DB	1021 AAGCTTCGGGTGAGCAATTCAGATTTCTTATCATGTCACCGGTCAAGGAATGGGCCAT 1080	
QY	1081 GGGGTTTGTGACCTTCTGGTGCACATCATGTTGCTGTATCGATGTCAGGGAGCATTAGG 1140	
DB	1081 GGGGTTTGTGACCTTCTGGTGCACATCATGTTGCTGTATCGATGTCAGGGAGCATTAGG 1140	
QY	1141 GTTTCACAGTCAGCGATGACCGGATGCTGTGTGCACTTGTGTGTCACCACTGATCAACTA 1200	
DB	1141 GTTTCACAGTCAGCGATGACCGGATGCTGTGTGCACTTGTGTGTCACCACTGATCAACTA 1200	
QY	1201 GGGACGATAGAGCTCTATGAATGTTTCAAACTTCTTCCCTCACACTCTAGGATCATGTGACA 1260	
DB	1201 GGGACGATAGAGCTCTATGAATGTTTCAAACTTCTTCCCTCACACTCTAGGATCATGTGACA 1260	
QY	1261 AAGGTGGGAGAGAGGGGGCTCTCTAGTGAAGGTGGAATGCACTGTCAGTGGGAAT 1320	
DB	1261 AAGGTGGGAGAGAGGGGGCTCTCTAGTGAAGGTGGAATGCACTGTCAGTGGGAAT 1320	
QY	1321 AGTGGCGCATCGCTTGTGTAATGAATAAAGGTGCTTGGGTGGGAGGAGTGAATATGA 1380	
DB	1321 AGTGGCGCATCGCTTGTGTAATGAATAAAGGTGCTTGGGTGGGAGGAGTGAATATGA 1380	
QY	1381 GGGAAATGATGTTGGTGGGGGATGTTCTCTTTTAAAGGAGGACCATTTGATTAATGGAAGA 1440	
DB	1381 GGGAAATGATGTTGGTGGGGGATGTTCTCTTTTAAAGGAGGACCATTTGATTAATGGAAGA 1440	
QY	1441 CAATGACACAAAGGTGGTGGGACAGTTTAAAGCTCGAATGCTGTAGGGGTGCTCAAGG 1500	
DB	1441 CAATGACACAAAGGTGGTGGGACAGTTTAAAGCTCGAATGCTGTAGGGGTGCTCAAGG 1500	
QY	1501 TTAAGAGATCAGCATCAGGGAGGAAAGGAGGATTAAGATTTCTTCTACCTGCTGCTG 1560	
DB	1501 TTAAGAGATCAGCATCAGGGAGGAAAGGAGGATTAAGATTTCTTCTACCTGCTGCTG 1560	
QY	1561 GGTGATGGGACAAAGGTGATGCTCAAGCAAGGAGGGGAGGATTCAGCGCAGAGATGCT 1620	
DB	1561 GGTGATGGGACAAAGGTGATGCTCAAGCAAGGAGGGGAGGATTCAGCGCAGAGATGCT 1620	
QY	1621 GTTGTGACACATGCGGGGGGGGAATTTGAGAGTTGGGGTTGACCAAGTGAAGTTTATGGG 1680	
DB	1621 GTTGTGACACATGCGGGGGGGGAATTTGAGAGTTGGGGTTGACCAAGTGAAGTTTATGGG 1680	
QY	1681 TGACCCAGAGAGAGAGACCCCACTGATGGGGGAAAGAGGTGCCAACAGGTGGGGACCAAGGT 1740	
DB	1681 TGACCCAGAGAGAGAGACCCCACTGATGGGGGAAAGAGGTGCCAACAGGTGGGGACCAAGGT 1740	
QY	1741 GTCAGTGAATCCGCTGACATGTTATTTGGAAAGTTACGTCGGGAATGGTTGGGCTCGAG 1800	
DB	1741 GTCAGTGAATCCGCTGACATGTTATTTGGAAAGTTACGTCGGGAATGGTTGGGCTCGAG 1800	
QY	1801 TGATCTAGGCTGGCTGGGACATGCTGTGATCTCTTAAATTTCTCATTTCCCAATTTAAGT 1860	
DB	1801 TGATCTAGGCTGGCTGGGACATGCTGTGATCTCTTAAATTTCTCATTTCCCAATTTAAGT 1860	
QY	1861 TGAATTTTAAATTAATCAATCAATGATCTCCAAATCTCTCCAAATTTACCAAAATATAGAAT 1920	
DB	1861 TGAATTTTAAATTAATCAATCAATGATCTCCAAATCTCTCCAAATTTACCAAAATATAGAAT 1920	
QY	1921 ATTTAGATGAATATGTTGGTGAAGTTTGGGCTCGCTTTTGGTATGATATGTTTGTATATAA 1980	
DB	1921 ATTTAGATGAATATGTTGGTGAAGTTTGGGCTCGCTTTTGGTATGATATGTTTGTATATAA 1980	
QY	1981 AATAATTTCT 2040	
DB	1981 AATAATTTCT 2040	
QY	2041 ACTTTTTTGTAGTGTGCGCATTTATGACACAAAGAACTATATCCATTTCTTCTATAGTCT 2100	

Db 2041 ACTTTTATAGTAGTGCCACTTATAGCACAAAATATATCCATTTCTATAGTCT 2100  
 Qy 2101 TGAATCCGACATCTTATTTAGCCATTTCTCAAAATGGCCAAACTAGGAATTTA 2160  
 Db 2101 TGAATCCGACATCTTATTTAGCCATTTCTCAAAATGGCCAAACTAGGAATTTA 2160  
 Qy 2161 ATACATTTCTGCCATAACATATTTAGTGCATTTCTCAAAATGGCCAAACTAGGAATTTA 2220  
 Db 2161 ATACATTTCTGCCATAACATATTTAGTGCATTTCTCAAAATGGCCAAACTAGGAATTTA 2220  
 Qy 2221 AACTTTCTTTGAGATTCATTAATTAATGCTACATGCTATCTTTTATAGATTCATCA 2280  
 Db 2221 AACTTTCTTTGAGATTCATTAATTAATGCTACATGCTATCTTTTATAGATTCATCA 2280  
 Qy 2281 ATATGCTCTCATTAGCATCTTCAATTTAGGAACTTGAATTAAGCCGCTTAAATATAGAG 2340  
 Db 2281 ATATGCTCTCATTAGCATCTTCAATTTAGGAACTTGAATTAAGCCGCTTAAATATAGAG 2340  
 Qy 2341 CCAAGTGACGGATCCATTTAAAGTGATCTTAATTTCTTACTTCTTATCTTTGGTGGCT 2400  
 Db 2341 CCAAGTGACGGATCCATTTAAAGTGATCTTAATTTCTTACTTCTTATCTTTGGTGGCT 2400  
 Qy 2401 TATGTTTATATATGTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460  
 Db 2401 TATGTTTATATATGTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460  
 Qy 2461 ACATATATATGGAATATAGCTTCACAGTCTAGA 2493  
 Db 2461 ACATATATATGGAATATAGCTTCACAGTCTAGA 2493

RESULT 2

AAF80484  
 ID AAF80484 standard; DNA; 499 BP.  
 XX AAF80484;  
 XX AAF80484;  
 DT 29-JUN-2001 (first entry)  
 DE Nucleotide sequence of a fragment of the Esr2 gene promoter.  
 KW Promoter; Esr2 gene; albumen; Angiosperm seed; embryo development;  
 KW transgenic plant; seed starch; seed oil; food; ss.  
 OS Zea mays.  
 PN FR2799203-A1.  
 PD 06-APR-2001.  
 XX 01-OCT-1999; 99PR-00012305.  
 XX 01-OCT-1999; 99PR-00012305.  
 PR (BIOG-) BIOGEMMA SAS.  
 PA Bonello JF, Rogowsky P, Perez P;  
 PI WPI; 2001-275283/29.  
 XX New plant promoters providing albumen-specific expression, useful for  
 XX preparing transgenic plants with altered starch and oil contents.  
 PT Claim 3; Page 43; 47pp; French.  
 XX The present sequence represents a fragment of the promoter of the Esr2  
 CC gene of maize. The promoter provides expression of attached sequences  
 CC that is specific to the region of albumen surrounding the embryo in  
 CC Angiosperm seeds and active in the early stages of albumen development.  
 CC The promoter provides specific expression of genes at an early stage of  
 CC embryo development, resulting in seeds of increased size, enriched in  
 CC starch and/or oil, or embryo-free seeds. The promoter is used to prepare

CC expression cassettes for preparation of transgenic Angiosperm plants with  
 CC improved agronomic or nutritional properties. Especially, these plants  
 CC produce seeds with altered starch and oil contents and are useful as  
 CC sources of human and animal foods  
 SQ Sequence 499 BP; 156 A; 90 C; 68 G; 185 T; 0 U; 0 Other;  
 Query Match 20.0%; Score 499; DB 4; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 5e-107;  
 Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1995 TTTTGTGCTTCCATATATGCTTAAATTTTATGTAGCAATGCCAATCTTTTATAGTAG 2054  
 Db 1 TTTTGTGCTTCCATATATGCTTAAATTTTATGTAGCAATGCCAATCTTTTATAGTAG 60  
 Qy 2055 TGTGCCACTATATAGCACAAAATATATCCATTTTCTAATAGTCTTGAATCCACATTC 2114  
 Db 61 TGTGCCACTATATAGCACAAAATATATCCATTTTCTAATAGTCTTGAATCCACATTC 120  
 Qy 2115 TATTTTAGCGATCTTCAAAATTTGGCACAAAATCTAGGAAAATTTAATACATCTTGCCA 2174  
 Db 121 TATTTTAGCGATCTTCAAAATTTGGCACAAAATCTAGGAAAATTTAATACATCTTGCCA 180  
 Qy 2175 TAACATATTTCTAGTGCATTTGTAATGCTCAATATAGCAAACTTCTTTTGTAA 2234  
 Db 181 TAACATATTTCTAGTGCATTTGTAATGCTCAATATAGCAAACTTCTTTTGTAA 240  
 Qy 2235 GATTCATTAATGCTACATCTCATCTTTTATAGAGTTCATCAATAGTCTCATTA 2294  
 Db 241 GATTCATTAATGCTACATCTCATCTTTTATAGAGTTCATCAATAGTCTCATTA 300  
 Qy 2295 GCATACCTTCAATTTAGGAATCTGATTAATAACCGCTTTAAATAGAGCCAGTGCAGATC 2354  
 Db 301 GCATACCTTCAATTTAGGAATCTGATTAATAACCGCTTTAAATAGAGCCAGTGCAGATC 360  
 Qy 2355 CATTTAAGGTGATCTTAATTTCTTACTTCTATCTTTGGTGGCTTATGTTATATAG 2414  
 Db 361 CATTTAAGGTGATCTTAATTTCTTACTTCTATCTTTGGTGGCTTATGTTATATAG 420  
 Qy 2415 TGTGGGTGGTGGATGATGTTCTTACACACCTACACACACGTTGGACATATATATGAA 2474  
 Db 421 TGTGGGTGGTGGATGATGTTCTTACACACCTACACACACGTTGGACATATATATGAA 480  
 Qy 2475 AATAGCTTCACAGTCTAGA 2493  
 Db 481 AATAGCTTCACAGTCTAGA 499

RESULT 3

AAF80482  
 ID AAF80482 standard; DNA; 1708 BP.  
 XX AAF80482;  
 XX AAF80482;  
 DT 29-JUN-2001 (first entry)  
 DE Nucleotide sequence of the promoter of the Esr3 gene of maize.  
 KW Promoter; Esr3 gene; albumen; Angiosperm seed; embryo development;  
 KW transgenic plant; seed starch; seed oil; food; ss.  
 OS Zea mays.  
 PN FR2799203-A1.  
 PD 06-APR-2001.  
 XX 01-OCT-1999; 99PR-00012305.  
 XX 01-OCT-1999; 99PR-00012305.  
 PR (BIOG-) BIOGEMMA SAS.













AB210100/c  
ID AB210100 standard; DNA; 8056 BP.  
XX  
AC AB210100,  
XX  
DT 16-JAN-2003 (first entry)  
XX  
DB Haematopoietic cell proliferation disorder related DNA sequence #240.  
XX  
XX Human; haematopoietic cell proliferation disorder; cytostatic;  
KM gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;  
KM cytosine methylation state; gene; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO20027272-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 26-MAR-2002; 2002WO-EP003401.  
XX  
XX 26-MAR-2001; 2001US-0278333P.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
XX  
XX Berlin K, Braun A, Dietler J, Gueig D, Howe A, Mueller J;  
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Iou E;  
PI Lewin A, Lipscher E, Maier S, Model P, Mueller V, Otto T, Pelet C;  
PI Schwobe I, Ziebarth H;  
XX  
DR WPI; 2003-018942/01.  
XX  
XX Detecting and differentiating between hematopoietic cell proliferative  
PT disorders, comprises contacting a target nucleic acid with a reagent that  
PT distinguishes between methylated and non-methylated CpG dinucleotides.  
XX  
XX Claim 28; SEQ ID NO 240; 117pp; English.  
XX  
XX The present invention describes a method for detecting and  
CC differentiating between hematopoietic cell proliferative disorders  
CC associated with at least 1 gene and/or their regulatory regions in a  
CC subject. The method comprises contacting a target nucleic acid in a  
CC biological sample obtained from the subject with at least 1 reagent,  
CC which distinguishes between methylated and non-methylated CpG  
CC dinucleotides within the target nucleic acid. AB209661 to AB211118  
CC represent specifically claimed nucleotide sequences from the present  
CC invention. Oligonucleotides from the present invention can be used: for  
CC differentiating between healthy hematopoietic cells and proliferative  
CC disorder hematopoietic cells; for differentiating between acute  
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for  
CC determining the cytosine methylation state and/or single nucleotide  
CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder  
CC related sequences and their complements; and as primers for the  
CC amplification of hematopoietic cell proliferation disorder related DNA  
CC sequences. The nucleotide sequences from the present invention can also  
CC be used for detecting a predisposition to, differentiation between  
CC subclases, diagnosis, prognosis, treatment and/or monitoring of  
CC haematopoietic cell proliferative disorders. The present method enables a  
CC highly specific classification of haematopoietic cell proliferative  
CC disorders allowing for improved and informed treatment of patients  
XX  
SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;  
Query Match 2.4%; Score 58.8; DB 8; Length 8056;  
Best Local Similarity 45.1%; Pred. No. 0.0036;  
Matches 219; Conservative 0; Mismatches 267; Indels 0; Gaps 0;  
QY 1852 AATTAAAGTTGAATTTTAAATCAATCAATGACTCCAAATCTCTCCAAATACCAA 1911  
DB 2361 AATTATATTAAATAAATAAATTTTAAATACGAAATTAATAAATTAATTTAA 2302  
QY 1912 ATATAGAAATTTAGATGAATATGTGGTGAGTGTGGGTCGGCTTGTGTAGTAGT 1971

DB 2301 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2242  
QY 1972 TTGTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2031  
DB 2241 TTTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2182  
QY 2032 GCATGCCCAACTTTTTTTAGTAGTGCCACTTATAGCACAATAAATAAATAA 2091  
DB 2181 ATAAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2122  
QY 2092 AATAGTCCTTGAATCCACATCTTATTTAGCCATTCITCAAAATTCGCAAACTAG 2151  
DB 2121 AATTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2062  
QY 2152 GAAATTTAATACATTTCTGCCATAACATATCTAGTGCAATATGTTAACTAGATGCTCA 2211  
DB 2061 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2002  
QY 2212 ATATTAGCAAACTCTTTTGTAGATTCATTAATATGTCACATCTTTTATAGA 2271  
DB 2001 AAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1942  
QY 2272 AGTTCATCAATAATGCTCATAGCATACCTCATTTTGGCACTTGATTAACCGCTT 2331  
DB 1941 ATATTATATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1882  
QY 2332 AATAATA 2337  
DB 1881 AATAATA 1876  
RESULT 12  
ADA71918  
ID ADA71938 standard; DNA; 2000 BP.  
XX  
AC ADA71938;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Rice gene, SEQ ID 5263.  
XX  
KW Plant; bacterial infection; fungal infection; viral infection; rice;  
KW gene; ds.  
XX  
OS Oryza sativa.  
XX  
FN WO2003000898-A1.  
XX  
PD 03-JAN-2003.  
XX  
PF 22-JUN-2001; 2001WO-IB001105.  
XX  
PR 22-JUN-2001; 2001WO-IB001105.  
XX  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri P, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX  
DR WPI; 2003-175290/17.  
XX  
XX Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.  
XX  
PS Claim 27; SEQ ID NO 5263; 899pp; English.  
XX  
CC The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that





XX	SQ	Sequence	6863 BP; 1997 A; 150 C; 1362 G; 3354 T; 0 U; 0 Other;
		Query Match	2.1%; Score 53.2; DB 4; Length 6863;
		Best Local Similarity	47.6%; Pred. No. 0.071;
		Matches 157; Conservative	0; Mismatches 173; Indels 0; Gaps 0;
QY	1840	TTCTCATTTCCCAATTTAAGTTGAATTTTAAATTCBAATCBAATGACTCCAAATCTCTCC	1899
DB	6224	TTCAAAATTTCCAACTTACCAAAACATATTAACCTCCCAACAACTAATAAAC	6165
QY	1900	AAATTTACCAAAATATAGAAATATTAGATGAATATGTTGGTGGAGTTTGGGCTCCGCTTT	1959
DB	6164	CAAAACACATTAACGAAAAATTTATATTTTACTAATAATAAATCATATATTTCTCATATT	6105
QY	1960	TGGTTAGTATGTTTGATATAAAAATAATTTCTCTCTTTTGGTCACATTCCAATATGTACTTA	2019
DB	6104	TTTTTCAAAATATTCAAACCAATTATTCAAATACCTTTCTACTTAAAAACCTACTTTTA	6045
QY	2020	AATTTTATATGTAGCAATGCCAACTTTTTTTTAGTAGTGTGCCACTTATAGCACAAAACTTA	2079
DB	6044	AATAATTAACAAAAAATTTTAATATTTTCTATCAAACTCCAACTTTTAAAAAAAATCA	5985
QY	2080	TATCCATTTTCTAATAGTCCTTGAAATCCACAATTTCTATTTTATAGCCATTCTTCAAAATG	2139
DB	5984	AATATATTAATTTATATATACACATTTCACTATATTTTCCAAAGCTCTTATACACCTAAT	5925
QY	2140	GCACAAAACCTAGAAAAATTTTAATCAATCTT	2169
DB	5924	ATTCAATTTTATAAACATCTAATTCGTTCT	5895

RESULT 14	
ABK31523/c	
ID	ABK31523 standard; DNA; 12669 BP.
XX	
XX	ABK31523;
XX	
XX	23-APR-2002 (first entry)
XX	
DE	Signal transduction associated gene modified complementary DNA #183.
DE	
XX	Human; signal transduction associated gene; cytosine methylation state;
XX	CpG island; signal transduction associated disease; solid tumour; cancer;
KW	antitumour; cytostatic; mutant; ds.
KW	

XX Oligonucleotide for diagnosis and therapy of diseases associated with  
XX signal transduction e.g. cancer, comprises chemically modified genomic  
XX sequences of genes associated with signal transduction.  
XX  
XX Claim 1; SEQ ID NO 366; 34pp; English.  
XX  
XX The present invention relates to chemically modified DNA sequences of  
XX signal transduction associated genes. The DNA sequences are chemically  
XX modified using a solution of bisulphite, hydrogen sulphite or disulphite.

Also disclosed are oligonucleotides and/or PNA oligomers for detecting the cytosine methylation state (CpG islands) of these genes, and a method for the diagnosis and/or therapy of genetic and epigenetic parameters of genes associated with signal transduction. The genomic DNA can be obtained from cells or cellular components which contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine, cerebellar-spinal fluid, tissue embedded in paraffin such as tissue from eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver, histologic object slides, and all their possible combinations. The sequences of the invention are useful for the diagnosis and therapy of diseases associated with signal transduction e.g. solid tumours and cancer. ARK31158-ARK31545 represent chemically pretreated genomic DNA sequences of different genes associated with signal transduction, or their complementary sequences. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

RESULT 15  
ABL70490/c  
ID ABL70490 standard: DNA; 12669 bp.

XX	ABL70490;
AC	
XX	
XX	01-JUL-2002 (first entry)
DT	
XX	
XX	Chemically treated cell signalling DNA sequence complementary to#190.
DE	
XX	
XX	Cell signalling; cytosine methylation; cell signalling disease; cancer
KW	tumour; cytostatic; ds.
KW	
XX	

OS Unidentified.

XX  
PN  
WO200202807-A2.

XX  
PD 10-JAN-2002.

XX  
PF 29-JUN-2001; 2001WO-EP007471.

PR 30-JUN-2000: 2000DE-01032529. XX

PR 01-SEP-2000; 2000DE-01043826.

XX  
PA (EPIG-) EPIGENOMICS AG.

PI Olek A. Pienenbrock C. Berlin K:

DR WPI; 2002-154758/20.

XX Nucleic acid, useful for diagnosis and therapy of diseases associated  
PT with cell signaling e.g. cancer, comprises chemically modified genomic  
PT sequences of genes associated with cell signaling.

PS Claim 1; SEQ ID NO 380; 24pp + Sequence Listing; English.

The invention relates to a nucleic acid comprising a sequence of at least 18 bases of a segment of chemically pretreated DNA of genes associated with cell signalling. The activity of the modified sequences of the invention may be described as cytostatic. The object of the invention is to provide the chemically modified DNA of genes associated with cell signalling, as well as oligonucleotides and/or DNA-oligomers for detecting cytosine methylations, as well as a method which is particularly suitable for the diagnosis and/or therapy of genetic and epigenetic parameters of genes associated with cell signalling. The chemically modified DNA provided by the invention is useful for diagnosis and therapy of diseases such as solid tumours and cancer. The sequences given in records AB170111-AB170636 represent chemically pretreated genomic DNA's of genes associated with cell signalling. Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office

Sequence 12669 BP: 3610 A: 224 C: 2687 G: 6148 T: 0 U: 0 Other: 0

Query Match	2.11	Score 52	DB 6	Length 12669
-------------	------	----------	------	--------------

Best Local Similarity 45.3%; Pred. No. 0.17;  
Matches 230; Conservative 0; Mismatches 275; Indels 3; Gaps 1;

Qy 143 ATTTTATCAAAATTTCTCATATAACAATTGAAACAATTTCTCTTTTGAAGAAGTGTAGAG 202

3300 ATAATCCCAACACTTTTAAAAAATAAACTAAATCTAATAAATCTAAATCTGAAAAATTAAGAA 3341

Qy 203 CACTAGCAACTGTCTACTAAAAGGTTCCCAATTTCTGGGTATAACAATCGCATGGTAA 262

3240 CCAGCTTACCCACACCGGAAAACTCTAAAAATACAAAAATTATCTAAACGTAAATAATAA 3181

QY 263 ATAAACAAGGAAATCCCTACTAAGAGCAGTAATTTGGCTAAACAATAGTGAACATTTT 322

Db 3180 AGCCTATATGCCCATATCCCACTACTTATAAATACTATACAAAA---ATTACTTA 3124

**Qv** 323 RANGTANATAGGGAATAGGAGCATCAATACTTGCTGTCTTTCTTCAGGGTGGTTCAGACGTCACCA 382

**d**

383 AAAGTGTGCCCCCCCCGTGGGCGAGTGGCA CACTGAAAATCTA CTGGTAATACATAAGAAAC 442

d7

d6

443 ATGGGCAAAATATAGAAACATATCTCAATTATGCAAAAGGTTCAATGGTCCATAAT 502

Db 3003 ATCAAAAAACTTCCAAAAAAATTAATAAAATTCACAAATCTATTTTCTACCTTATA 2844

QV 503 TATTGTAGACATTTTGAATTTATTTTTCAGCCAAACCCTTTTAATTGCGTTAAAAAGA 562

27  
Db  
2943 AATTTACCGATTCTTAAANNTTTCATATATAAATAAATCATACATATATATAAATTTTA 2884

.....

Qy 563 GTTAGATATTAAATATTTATTCAGTTTATAGTTATTATGGGACATTTATTTACTTAACATA 622

Db 2883 TATCTAAATTTTTCATTTTACATAAATATTTTAAAAATTTATCCCATTTTATTATTATTATA 2824

Qy 623 ACTTCTAGGGTTTTAAAGTAAATTTTG 650

2823 ATTATTTTAAAAATAAAATCTTCG 279

Search completed: April 9, 2005, 02:57:34  
Job time : 1158 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2005, 23:00:44 ; Search time 10359 seconds  
(without alignments)  
11661.247 Million cell updates/sec

Title: US-10-089-612-2

Perfect score: 2493

Sequence: 1 agctttccggtgatgaag.....aaatagcttcacagtctaga 2493

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.cm.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.ste.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2493	100.0	2493	6	AX111989 Sequence
2	2493	100.0	3542	6	AJ251320 Zea mays
3	499	20.0	499	6	AX111992 Sequence
4	341	13.7	1708	6	AX111990 Sequence
5	341	13.7	1708	8	ZMA251319
6	339	13.6	507	6	AX111993 Sequence
7	236.8	9.5	1232	6	AX111991 Sequence
8	216.2	8.7	265	6	AX111994 Sequence
9	163.6	6.6	531	6	AX111988 Sequence
10	163.6	6.6	531	8	ZMA251118
11	77.6	3.1	203163	8	AF546188 Contiguou
12	76.4	3.1	81372	2	AC139056 Mus muscu
13	76.4	3.1	105035	2	AC145452 Zea mays
14	74	3.0	7218	6	166494 Sequence 14
15	72.6	2.9	192816	2	AC148162 Zea mays
16	70.4	2.8	178657	2	AC148234 Zea mays
17	70	2.8	211974	2	AC150184 Zea mays
18	67.2	2.7	149327	2	AC146976 Zea mays
19	67.2	2.7	288479	2	AC146814 Zea mays

C 20	66.8	2.7	129466	2	AC148099	Zea mays
C 21	63.8	2.6	181249	2	AC148479	Zea mays
C 22	63.6	2.6	10856	8	AY146813	Zea mays
C 23	62.4	2.5	10858	8	AY146787	Zea mays
C 24	62.4	2.5	186998	2	CR388057	Danio rer
C 25	62	2.5	993	8	AY232485	Zea mays
C 26	62	2.5	8056	6	AX599046	Sequence
C 27	62	2.5	10839	8	AY146796	Zea mays
C 28	62	2.5	10833	8	AY146816	Zea mays
C 29	62	2.5	10854	8	AY146798	Zea mays
C 30	62	2.5	10854	8	AY146814	Zea mays
C 31	62	2.5	10854	8	AY146815	Zea mays
C 32	62	2.5	10855	8	AY146800	Zea mays
C 33	62	2.5	10856	8	AY146791	Zea mays
C 34	62	2.5	10856	8	AY146797	Zea mays
C 35	62	2.5	10856	8	AY146806	Zea mays
C 36	62	2.5	10858	8	AY146786	Zea mays
C 37	62	2.5	10858	8	AY146788	Zea mays
C 38	62	2.5	10858	8	AY146789	Zea mays
C 39	62	2.5	10858	8	AY146790	Zea mays
C 40	62	2.5	10858	8	AY146792	Zea mays
C 41	62	2.5	10858	8	AY146793	Zea mays
C 42	62	2.5	10858	8	AY146794	Zea mays
C 43	62	2.5	10858	8	AY146795	Zea mays
C 44	62	2.5	10858	8	AY146799	Zea mays
C 45	62	2.5	10858	8	AY146801	Zea mays

## ALIGNMENTS

RESULT 1	AX111989	2493 bp	DNA	linear	PAT 01-MAY-2001
LOCUS	Sequence 2 from Patent WO0125439.				
DEFINITION	AX111989				
ACCESSION	AX111989.1	GI:13938897			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD				
	clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE					
AUTHORS	Bonello, J.P., Rogowsky, P. and Perez, P.				
TITLE	Plant seed endosperm-specific promoter				
JOURNAL	Patent: WO 0125439-A 2 12-APR-2001;				
	Biogemma (FR)				

## FEATURES

source	Location/Qualifiers
1..2493	/organism="Zea mays"
	/mol_type="unassigned DNA"
	/db_xref="taxon:4577"

## ORIGIN

Query Match	100.0%;	Score 2493;	DB 6;	Length 2493;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2493;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AAGCTTTTCGGTGATGAAGCCTGTATATCTTAACGAGATGCTGAAAACAATAGTTA	60	
DB	1	AAGCTTTTCGGTGATGAAGCCTGTATATCTTAACGAGATGCTGAAAACAATAGTTA	60	
QY	61	GCTGTGTTTTTGAGGACCTTCGGAAGATGAAGCCCCCAACACATCCCATGATCAAGTC	120	
DB	61	GCTGTGTTTTTGAGGACCTTCGGAAGATGAAGCCCCCAACACATCCCATGATCAAGTC	120	
QY	121	CCCATGATTCGAAAAAGCAAAATTTTATCAAAATTTCTATATAAATCTTGAACAAATTT	180	
DB	121	CCCATGATTCGAAAAAGCAAAATTTTATCAAAATTTCTATATAAATCTTGAACAAATTT	180	
QY	181	TCTCTCTTTTCAAAAGTGTAGACCACTGTCTACTAAAAAGSTTCCCAAAATTTCT	240	

Db 181 TCTCTTTTGAAGGTGTAGAGCACTAGCAACTGTCTACTTAAAGGTTCCCAATTTCT 240  
Qy 241 GGGTATAACAATTCGATGGTAAATAACAACAAGGAAATCTTACTAAGACGATPAATTTGG 300  
Db 241 GGGTATAACAATTCGATGGTAAATAACAACAAGGAAATCTTACTAAGACGATPAATTTGG 300  
Qy 301 CTAACCAATAGTAGAGCAATTTAAATGTAATAGGGAATAGAGCATGCAATCTTGTGTTTC 360  
Db 301 CTAACCAATAGTAGAGCAATTTAAATGTAATAGGGAATAGAGCATGCAATCTTGTGTTTC 360  
Qy 361 TTTTCAGGGTTTGTATGTCCTCAAAAAGTGTGCCCTCGGGGCAAGTTCGCAACACTCAAAAT 420  
Db 361 TTTTCAGGGTTTGTATGTCCTCAAAAAGTGTGCCCTCGGGGCAAGTTCGCAACACTCAAAAT 420  
Qy 421 CTACTCGTATACATAAAGAAACATGCGCACAAAATAGAAACAACTACTCAAAATATGAAA 480  
Db 421 CTACTCGTATACATAAAGAAACATGCGCACAAAATAGAAACAACTACTCAAAATATGAAA 480  
Qy 481 AAGGTTCAAAATGGTCTCTAATAATTTATGTAGACATTTTAGATTTTATAGACCAAAACC 540  
Db 481 AAGGTTCAAAATGGTCTCTAATAATTTATGTAGACATTTTATAGATTTTATAGACCAAAACC 540  
Qy 541 ATTAAATTTGGTTTAAATGAGTTAGATATTAATATTTATTCAGTTTATAGTTATTTGG 600  
Db 541 ATTAAATTTGGTTTAAATGAGTTAGATATTAATATTTATTCAGTTTATAGTTATTTGG 600  
Qy 601 GACATTTATTTACTTAACATACTTCTAGGGTTTAAAGTAAATTTGGGTCCTAGT 660  
Db 601 GACATTTATTTACTTAACATACTTCTAGGGTTTAAAGTAAATTTGGGTCCTAGT 660  
Qy 661 TGGAACTAGCTCAGATTTGCTTGAATTTCCATAAAGTCCAGGTTCTTTTAGCAAAAAT 720  
Db 661 TGGAACTAGCTCAGATTTGCTTGAATTTCCATAAAGTCCAGGTTCTTTTAGCAAAAAT 720  
Qy 721 CCAGGTGAAAGGGGGAGATAGGTGTGACCGATATCTCTAAATTTGATGCTGTGAC 780  
Db 721 CCAGGTGAAAGGGGGAGATAGGTGTGACCGATATCTCTAAATTTGATGCTGTGAC 780  
Qy 781 GGCACTAGGATGCTCAGATTTAAATGGTGAATGTGCAAGCAAGCGCACACAGATGGAGA 840  
Db 781 GGCACTAGGATGCTCAGATTTAAATGGTGAATGTGCAAGCAAGCGCACACAGATGGAGA 840  
Qy 841 ATGGCTTACGACGGTGGCTACTAGAGCTGGCTAGCTGACCAATGGAGGCTCGGTCA 900  
Db 841 ATGGCTTACGACGGTGGCTACTAGAGCTGGCTAGCTGACCAATGGAGGCTCGGTCA 900  
Qy 901 AGGTCAAAAATTTGTTGCCAAGCACTGTGGCTCACTGAGTCGATTCGATGACACATATCAA 960  
Db 901 AGGTCAAAAATTTGTTGCCAAGCACTGTGGCTCACTGAGTCGATTCGATGACACATATCAA 960  
Qy 961 GGTGAGGCTCAACAGAGGGGCAAGATGATGATGGTGGTGTCTCGATGGAGGGG 1020  
Db 961 GGTGAGGCTCAACAGAGGGGCAAGATGATGATGGTGGTGTCTCGATGGAGGGG 1020  
Qy 1021 AAATCTCGGTGAGCAATTTCAAGATTTCTATCATGTGACCGGGTCAGGGAATGGGCGAT 1080  
Db 1021 AAATCTCGGTGAGCAATTTCAAGATTTCTATCATGTGACCGGGTCAGGGAATGGGCGAT 1080  
Qy 1081 GGGGTTTGTACCTTCTGTGACATCATGTTGCTGTATTCGATGTCAAGGAGCAATTAGG 1140  
Db 1081 GGGGTTTGTACCTTCTGTGACATCATGTTGCTGTATTCGATGTCAAGGAGCAATTAGG 1140  
Qy 1141 GTTCACGATCAGCGATGACGGGCGATGGTGGCACTTGTGTCAACATGGTTCGATCAACTA 1200  
Db 1141 GTTCACGATCAGCGATGACGGGCGATGGTGGCACTTGTGTCAACATGGTTCGATCAACTA 1200  
Qy 1201 GGGAGATAGACTCTATGAGTTTCAACACTTCTCACTCTAGGGATCATGTGTACA 1260  
Db 1201 GGGAGATAGAGCTCTATGAGTTTCAACACTTCTCACTCTAGGGATCATGTGTACA 1260  
Qy 1261 AAGGTGGGGAGGACGGGGGCTCTCTAGTGGGGTGGAAATGCAAGTTCTGTCAAGTGGGAAT 1320  
Db 1261 AAGGTGGGGAGGACGGGGGCTCTCTAGTGGGGTGGAAATGCAAGTTCTGTCAAGTGGGAAT 1320

Qy 1321 AGTGGCGGCATCGCTGTGTAAATGAATAAAGTCTTTGGTGGCTGGGAAGTGCATATGA 1380  
Db 1321 AGTGGCGGCATCGCTGTGTAAATGAATAAAGTCTTTGGTGGCTGGGAAGTGCATATGA 1380  
Qy 1381 GGGAGATAGTTGGTGGCGGGAATGTTCTTTTATAAGGAGACCAATTTAATATGAGAGA 1440  
Db 1381 GGGAGATAGTTGGTGGCGGGAATGTTCTTTTATAAGGAGACCAATTTAATATGAGAGA 1440  
Qy 1441 CAATGACACAAGAGGTGGTGGCGACAGTTTAAAGTCGAAATGCTGTAGGGTGTCTCAAGG 1500  
Db 1441 CAATGACACAAGAGGTGGTGGCGACAGTTTAAAGTCGAAATGCTGTAGGGTGTCTCAAGG 1500  
Qy 1501 TTAAGAATCAGGCATCAGGAGGAAAGCGAGGATAAAATTTCTTTACTCCAGTTGTGG 1560  
Db 1501 TTAAGAATCAGGCATCAGGAGGAAAGCGAGGATAAAATTTCTTTACTCCAGTTGTGG 1560  
Qy 1561 GGTGATGGGGAACAAGGTAGTGTCTCAAGCAAGGAGGCGAGTTTCAGGCGAGAGATGGCT 1620  
Db 1561 GGTGATGGGGAACAAGGTAGTGTCTCAAGCAAGGAGGCGAGTTTCAGGCGAGAGATGGCT 1620  
Qy 1621 GTTGTGACACATCGGGGGGGGGAATTCGAGGTTTGGGTTGACCAAGTCAAGTTATGGGG 1680  
Db 1621 GTTGTGACACATCGGGGGGGGGAATTCGAGGTTTGGGTTGACCAAGTCAAGTTATGGGG 1680  
Qy 1681 TGACCCAGAGAGAGACCCACTGATGGGGAATAAAGTGGCGCAAGTGGGGAGCCAGGT 1740  
Db 1681 TGACCCAGAGAGAGACCCACTGATGGGGAATAAAGTGGCGCAAGTGGGGAGCCAGGT 1740  
Qy 1741 GTCAAGTCACTCACTGTGACATGTTAATGGAAAGTACGTCCGGAAATGCTTTGGGGCTCGAG 1800  
Db 1741 GTCAAGTCACTCACTGTGACATGTTAATGGAAAGTACGTCCGGAAATGCTTTGGGGCTCGAG 1800  
Qy 1801 TGAATCTAGGCTGGGCTGGGCACTGTGCTGATCTTTAATTTCTCATTTCCCAATTTAAGT 1860  
Db 1801 TGAATCTAGGCTGGGCTGGGCACTGTGCTGATCTTTAATTTCTCATTTCCCAATTTAAGT 1860  
Qy 1861 TGAATTTTAAATCAAAATCAAAATCAAAATCTCTCCAAAATTTACAAAATATAGAAAT 1920  
Db 1861 TGAATTTTAAATCAAAATCAAAATCAAAATCTCTCCAAAATTTACAAAATATAGAAAT 1920  
Qy 1921 ATTATAGTGAATATGTTGGTGGAGTTGGGCTGGCTTTGGTTAGTATGTTGTATATAA 1980  
Db 1921 ATTATAGTGAATATGTTGGTGGAGTTGGGCTGGCTTTGGTTAGTATGTTGTATATAA 1980  
Qy 1981 AATAATTTCTCTCTTTGTCTCACTTCCAAATATTTGACTTAAATTTTATGTAGCAATGCCA 2040  
Db 1981 AATAATTTCTCTCTTTGTCTCACTTCCAAATATTTGACTTAAATTTTATGTAGCAATGCCA 2040  
Qy 2041 ACTTTTATAGTGTGCGACCTTATAGCACAAAAATCTATATCAATTTTCTAAATAGTCT 2100  
Db 2041 ACTTTTATAGTGTGCGACCTTATAGCACAAAAATCTATATCAATTTTCTAAATAGTCT 2100  
Qy 2101 TGAATTTCAACATTTCTATTTTATAGCCATTTCTCAAAATTTGGCAAAAATAGGAAATTTA 2160  
Db 2101 TGAATTTCAACATTTCTATTTTATAGCCATTTCTCAAAATTTGGCAAAAATAGGAAATTTA 2160  
Qy 2161 ATACATCTTGGCAATATATTTCTAGTGCAAATGTTAACTAGATTCCTCATATATAGCA 2220  
Db 2161 ATACATCTTGGCAATATATTTCTAGTGCAAATGTTAACTAGATTCCTCATATATAGCA 2220  
Qy 2221 AACTCTTTTGTAGATTCATTAATATTTGCTACATATGCAATCTTTTATAGAGTTTCATCA 2280  
Db 2221 AACTCTTTTGTAGATTCATTAATATTTGCTACATATGCAATCTTTTATAGAGTTTCATCA 2280  
Qy 2281 ATAAATGCTCTAGTACATCTTCAATTTTGGAACTTTGATTAAGCCCTTTAAATATAGAG 2340  
Db 2281 ATAAATGCTCTAGTACATCTTCAATTTTGGAACTTTGATTAAGCCCTTTAAATATAGAG 2340  
Qy 2341 CCAAGTGCAGGATCCATTTTAAAGGTGATTTCTTAAATTTCTTACTTCTTATCTTGGTGGCT 2400  
Db 2341 CCAAGTGCAGGATCCATTTTAAAGGTGATTTCTTAAATTTCTTACTTCTTATCTTGGTGGCT 2400

Qy 2401 TATGTTTATATATGTTGGTGGTTCGAATGATGTTCTACACACTACACACAGTGG 2460  
Db 2401 TATGTTTATATATGTTGGTGGTTCGAATGATGTTCTACACACTACACACAGTGG 2460  
Qy 2461 ACATATATATGGAATAGCTTCACAGTCTAGA 2493  
Db 2461 ACATATATATGGAATAGCTTCACAGTCTAGA 2493

RESULT 2  
ZMA251320  
LOCUS Zea mays Ear2 gene, 5' flanking region. linear PLN 26-APR-2000  
DEFINITION A251320  
ACCESSION A251320  
VERSION A251320.1 GI:7657939  
KEYWORDS 5' flanking region; Ear2 gene.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1  
AUTHORS Bonello,J.F., Opsahl-Perstad,H.G., Perez,P., Dumas,C. and Rogowsky,P.M.  
TITLE Ear genes show different levels of expression in the same region of maize endosperm  
JOURNAL Gene 246 (1-2), 219-227 (2000)  
MEDLINE 20231747  
PUBMED 10767543  
REFERENCE 2 (bases 1 to 3542)  
AUTHORS Rogowsky,P.M.  
TITLE Direct Submission  
JOURNAL Submitted (24-NOV-1999) Rogowsky P.M., Reproduction et Developpement des Plantes, ENS-Lyon, 46 allée d'Italie, P-69364 Lyon Cedex 07, FRANCE

FEATURES  
source  
1..3542  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/variety="hybrid DHSxDH7"  
/db\_xref="taxon:4577"  
/map="chromosome arm 18"  
/tissue type="leaf"  
/dev stage="seedling"  
misc\_feature  
41..505  
/note="homology to non-LTR of several maize retrotransposons"  
LTR  
506..1050  
/note="homology to several maize retrotransposons"  
gene  
3047..3542  
/gene="Ear2"  
promoter  
3047..3542  
/gene="Ear2"  
/note="homology to Ear1 and Ear3 promoters"

ORIGIN  
Query Match 100.0%; Score 2493; DB 8; Length 3542;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2493; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAGCTTTCCGGTGATGAGCACCCTGTAATCTTAACAGCATGCTGGAACAATAGTTA 60  
Db 1050 AAGCTTTCCGGTGATGAGCACCCTGTAATCTTAACAGCATGCTGGAACAATAGTTA 1109  
Qy 61 GCTGTGTTTTCAGGACCTTCGGAAGATGAAGGCCCCCAACACATCCCATGCATCAAGTC 120  
Db 1110 GCTGTGTTTTCAGGACCTTCGGAAGATGAAGGCCCCCAACACATCCCATGCATCAAGTC 1169  
Qy 121 CCATGATGCAAAAAGCAAAATTTTATCAAAATTTCTATATAAAACACTTGAACAAT 180  
Db 1170 CCATGATGCAAAAAGCAAAATTTTATCAAAATTTCTATATAAAACACTTGAACAAT 1229  
Qy 181 TCTCTTTTGAAGAGTGTAGACCACTAGCAACTGCTCTACTATAAAGGTTCCCAATTTCT 240

Db 1230 TCTCTTTTGAAGAGTGTAGACCACTAGCAACTGCTCTACTATAAAGGTTCCCAATTTCT 1289  
Qy 241 GGGTATAACAATCGCATGGTAAATAACACAAAGAAATCTCTACTAAGAGCAGTAATTTGG 300  
Db 1290 GGGTATAACAATCGCATGGTAAATAACACAAAGAAATCTCTACTAAGAGCAGTAATTTGG 1349  
Qy 301 CTAATAACAATAGTGAAGCAATTTAAATGTAATAGGAATAGGAGCATGCAATCTTGTGTC 360  
Db 1350 CTAATAACAATAGTGAAGCAATTTAAATGTAATAGGAATAGGAGCATGCAATCTTGTGTC 1409  
Qy 361 TTTTCAGGGTTTTCATGCTCTCAAAAGTGTGCCCCCTGGGGCAGTTGCAACATCAAAAT 420  
Db 1410 TTTTCAGGGTTTTCATGCTCTCAAAAGTGTGCCCCCTGGGGCAGTTGCAACATCAAAAT 1469  
Qy 421 CTACTCTATATACATAAAGAAACATGGGCACAAATTAAGAAACAATACTCAAAATATGAAA 480  
Db 1470 CTACTCTATATACATAAAGAAACATGGGCACAAATTAAGAAACAATACTCAAAATATGAAA 1529  
Qy 481 AAGGTTCAAAATGCTCTATTAATTTATTTAGACATTTTATAGAAATTTTATAGCAACCAACC 540  
Db 1530 AAGGTTCAAAATGCTCTATTAATTTATTTAGACATTTTATAGAAATTTTATAGCAACCAACC 1589  
Qy 541 ATTTAAATTTGGTTTAAATAGAGTAGATATTAATTTATTTCAAGTTTATAGTTATTTGG 600  
Db 1590 ATTTAAATTTGGTTTAAATAGAGTAGATATTAATTTATTTCAAGTTTATAGTTATTTGG 1649  
Qy 601 GACATTTATTTACTTAACTATATACTTTCTAGGGTTTAAAGTAATAATTTTGGTCCCTAGT 660  
Db 1650 GACATTTATTTACTTAACTATATACTTTCTAGGGTTTAAAGTAATAATTTTGGTCCCTAGT 1709  
Qy 661 TGGAACTAGCTCAGATGCTGGTTGATTTTCCATAAAGTCAGGTTCTTTAGCAAAAT 720  
Db 1710 TGGAACTAGCTCAGATGCTGGTTGATTTTCCATAAAGTCAGGTTCTTTAGCAAAAT 1769  
Qy 721 CCAGGTGAACAGGGGGAGATAGGTGTTGACCGATATCTCTAAATTTTGTGATCGTGAC 780  
Db 1770 CCAGGTGAACAGGGGGAGATAGGTGTTGACCGATATCTCTAAATTTTGTGATCGTGAC 1829  
Qy 781 GGCACATGATGCTCTCAGATTAATAATGTTGGATGTGCAAGCGACGCGCACACATGAGGA 840  
Db 1830 GGCACATGATGCTCTCAGATTAATAATGTTGGATGTGCAAGCGACGCGCACACATGAGGA 1889  
Qy 841 ATGGCTTCAGAGCGTGGGCTACTAGAGCTGGCTACGTCAACCAATGAGGGGCTCGGTCA 900  
Db 1890 ATGGCTTCAGAGCGTGGGCTACTAGAGCTGGCTACGTCAACCAATGAGGGGCTCGGTCA 1949  
Qy 901 AGGTCAAAATTTTGGCCAAAGCCACTGTGGCTCAGATGAGTGTGATGAGCACAATATCAA 960  
Db 1950 AGGTCAAAATTTTGGCCAAAGCCACTGTGGCTCAGATGAGTGTGATGAGCACAATATCAA 2009  
Qy 961 GGTGAGGGTCAACCAAGGGGGCAAGATCGATGTGCAATGTTGTTCTGATGGAAGGGG 1020  
Db 2010 GGTGAGGGTCAACCAAGGGGGCAAGATCGATGTGCAATGTTGTTCTGATGGAAGGGG 2069  
Qy 1021 AAACCTCGGTGAGCAATTTCAAGATTTCTATCATGTCAACCGGTCAGGGATGGCGCAT 1080  
Db 2070 AAACCTCGGTGAGCAATTTCAAGATTTCTATCATGTCAACCGGTCAGGGATGGCGCAT 2129  
Qy 1081 GGGGTTTGGTACCTTCTGGTGCACATCATGTTGCTGTATCGATGTCAGGGAGCATTTAGG 1140  
Db 2130 GGGGTTTGGTACCTTCTGGTGCACATCATGTTGCTGTATCGATGTCAGGGAGCATTTAGG 2189  
Qy 1141 GTTCAGAGTCAAGCGATGACGGGCACTGGTGGGACTGTGTGTCACATGTTTGGATCAACTA 1200  
Db 2190 GTTCAGAGTCAAGCGATGACGGGCACTGGTGGGACTGTGTGTCACATGTTTGGATCAACTA 2249  
Qy 1201 GGGACGATAGCTCTATGAAGTTTCAAACTTCTCTCACTCTAGGGATCATGGTGACA 1260  
Db 2250 GGGACGATAGCTCTATGAAGTTTCAAACTTCTCTCACTCTAGGGATCATGGTGACA 2309  
Qy 1261 AAGGTGGGAGGACGGGGCGTCTCTAGTGAGGGTGAATGTCAGTCTGTGACGTGGGAAT 1320

Db 2110 AAGTGGGGAGGACGGGGCGCTCTCTAGTGGGGTGGAAATGCAAGTCTCTGTCACTGGGAAT 2369  
 Qy 1321 AGTGGGGCATCGCTGTAAATGAATAAAGGTCTCTGGGTGGCTGGGAAGTCAATATGA 1380  
 Db 2370 AGTGGGGCATCGCTGTAAATGAATAAAGGTCTCTGGGTGGCTGGGAAGTCAATATGA 2429  
 Qy 1381 GGGAAATAGTGGTGGGGGATGTTCTTTTATAGGAGGACCACTTGAATTAATGGAAGA 1440  
 Db 2430 GGGAAATAGTGGTGGGGGATGTTCTTTTATAGGAGGACCACTTGAATTAATGGAAGA 2489  
 Qy 1441 CAATGACACAAAGGGTGGTGGACAGTTTAAAGCTCGAATGCTCTAGGGGTGCTCAAGG 1500  
 Db 2490 CAATGACACAAAGGGTGGTGGACAGTTTAAAGCTCGAATGCTCTAGGGGTGCTCAAGG 2549  
 Qy 1501 TTAAGAGATCAGGCATCAGGGAGGAAGGCAAGGATTAATAATTTCTTTTACTCAGTTGGG 1560  
 Db 2550 TTAAGAGATCAGGCATCAGGGAGGAAGGCAAGGATTAATAATTTCTTTTACTCAGTTGGG 2609  
 Qy 1561 GGTGATGGGACAAAGGTAGTGTCTCAAGCAAGGGAGGGCGAGTTTCAGCGCAGAGATGCT 1620  
 Db 2610 GGTGATGGGACAAAGGTAGTGTCTCAAGCAAGGGAGGGCGAGTTTCAGCGCAGAGATGCT 2669  
 Qy 1621 GTTGTGACACATCGGGGGGGGGAATTTGGAGTTGGGTGACCAAGTGAAGTATGCGG 1680  
 Db 2670 GTTGTGACACATCGGGGGGGGGAATTTGGAGTTGGGTGACCAAGTGAAGTATGCGG 2729  
 Qy 1681 TGACCCAGAGAGAGACCCACTGATGGGGAAGGAGGTCGCAAGGTGGGAGCAAGGT 1740  
 Db 2730 TGACCCAGAGAGAGACCCACTGATGGGGAAGGAGGTCGCAAGGTGGGAGCAAGGT 2789  
 Qy 1741 GTCAGTGAATCAGCGTGCACATGTTATGGAAGTACGTTCGGGAAGTGTGGGCTCGAG 1800  
 Db 2790 GTCAGTGAATCAGCGTGCACATGTTATGGAAGTACGTTCGGGAAGTGTGGGCTCGAG 2849  
 Qy 1801 TGATCTAGCTGGCTCGGGCACTGTGCTGATCTTAAATTTCTTCATCTCCATTTAAGT 1860  
 Db 2850 TGATCTAGCTGGCTCGGGCACTGTGCTGATCTTAAATTTCTTCATCTCCATTTAAGT 2909  
 Qy 1861 TGAATTTTAAATCAAAATCAATGACTCCAAATCTCTCCAAATTTACCAAAATATAGAT 1920  
 Db 2910 TGAATTTTAAATCAAAATCAATGACTCCAAATCTCTCCAAATTTACCAAAATATAGAT 2969  
 Qy 1921 ATTTAGATCAATATGTTGGTGGAGTTGGGCTCGGCTTTGGTTAGTATGTTGTATAAA 1980  
 Db 2970 ATTTAGATCAATATGTTGGTGGAGTTGGGCTCGGCTTTGGTTAGTATGTTGTATAAA 3029  
 Qy 1981 AATAATTTCTCTCTCTTTGTCATCTCCAAATTTAGCACTTAAATTTTATGTAGCAATGCA 2040  
 Db 3030 AATAATTTCTCTCTCTTTGTCATCTCCAAATTTAGCACTTAAATTTTATGTAGCAATGCA 3089  
 Qy 2041 ACTTTTTTTAGTGTGCGCATTTATAGCACAAAAACCTATATCCATTTCTTAATAGTCT 2100  
 Db 3090 ACTTTTTTTAGTGTGCGCATTTATAGCACAAAAACCTATATCCATTTCTTAATAGTCT 3149  
 Qy 2101 TGAATCCCACTCTCTTTTATGCGCAATTTTACCAATTTTACCAAACTAGGAAATTTA 2160  
 Db 3150 TGAATCCCACTCTCTTTTATGCGCAATTTTACCAATTTTACCAAACTAGGAAATTTA 3209  
 Qy 2161 ATACATCTCTGCGCATTAATTTAGTGGAAATGTTAACTAGATGCTCAATATTTAGCA 2220  
 Db 3210 ATACATCTCTGCGCATTAATTTAGTGGAAATGTTAACTAGATGCTCAATATTTAGCA 3269  
 Qy 2221 AACTCTCTTTTGTAGATTTCAATTAATTTGCTACATGCTATCTTTTATAGAGTTTATCA 2280  
 Db 3270 AACTCTCTTTTGTAGATTTCAATTAATTTGCTACATGCTATCTTTTATAGAGTTTATCA 3329  
 Qy 2281 ATAATGCTCTATTTAGCATCTCTTTTATAGGAACTTGAATTAAGGCGCTTAAATAGAG 2340  
 Db 3330 ATAATGCTCTATTTAGCATCTCTTTTATAGGAACTTGAATTAAGGCGCTTAAATAGAG 3389  
 Qy 2341 CCAAGTGAAGGATCCATTTAAAGGTGATTTCTTAATTTCTTACTTCTTACTTTTGGTGGCT 2400  
 Db 3390 CCAAGTGAAGGATCCATTTAAAGGTGATTTCTTAATTTCTTACTTCTTACTTTTGGTGGCT 3449

Qy 2401 TATGTTTATATATGTTGGGTGGTGAATGATGTTCTTACCACTACCACTACCACTGG 2460  
 Db 3450 TATGTTTATATATGTTGGGTGGTGAATGATGTTCTTACCACTACCACTACCACTGG 3509  
 Qy 2461 ACATATATATGGAATAATAGCTTTCACAGTCTTACA 2493  
 Db 3510 ACATATATATGGAATAATAGCTTTCACAGTCTTACA 3542

RESULT 3  
 AX111992  
 LOCUS Sequence 5 from Patent WO0125439. 499 bp DNA linear PAT 01-MAY-2001  
 DEFINITION  
 AX111992  
 ACCESSION  
 AX111992.1 GI:13938900  
 VERSION  
 KEYWORDS  
 SOURCES  
 Zea mays  
 Zea mays  
 ORGANISM  
 Zea mays  
 Sukuayota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
 1  
 AUTHORS Bonello, J.F., Rogovsky, P. and Perez, P.  
 TITLE Plant seed endosperm-specific promoter  
 JOURNAL Patent: WO 0125439-A 5 12-APR-2001;  
 BIOGEMMA (FR)  
 LOCATION/Qualifiers  
 source  
 1. .499  
 /organism="Zea mays"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4577"

ORIGIN  
 Query Match 20.0%; Score 499; DB 6; Length 499;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-94; Indels 0; Gaps 0;  
 Matches 499; Conservative 0; Mismatches 0;  
 Qy 1995 TTTTGTCACTTCCAAATATGACTTAAATTTTATGAGCAATGCAATCTTTTAGTAG 2054  
 Db 1 TTTTGTCACTTCCAAATATGACTTAAATTTTATGAGCAATGCAATCTTTTAGTAG 60  
 Qy 2055 TGTGCGCATTTAGCAAAATATATATCTTCTATATAGTCTTGAATCCACATTC 2114  
 Db 61 TGTGCGCATTTAGCAAAATATATCTTCTATATAGTCTTGAATCCACATTC 120  
 Qy 2115 TATTTTATAGCAATCTTCAAAATTTGGCAAAAATAGGAAATTTAATACTTTTGGCA 2174  
 Db 121 TATTTTATAGCAATCTTCAAAATTTGGCAAAAATAGGAAATTTAATACTTTTGGCA 180  
 Qy 2175 TACATATCTAGTGGCAAAATGTTTAACTAGATTTCTCAATATTTAGCAAACTTTTGTAA 2234  
 Db 181 TACATATCTAGTGGCAAAATGTTTAACTAGATTTCTCAATATTTAGCAAACTTTTGTAA 240  
 Qy 2235 GATTCATTAATTTAGTGTACATTTGCAATCTTTTAAAGTTCATCAATAATGCTCATTA 2294  
 Db 241 GATTCATTAATTTAGTGTACATTTGCAATCTTTTAAAGTTCATCAATAATGCTCATTA 300  
 Qy 2295 GCATCTTCAATTTAGGAACTTGAATTAAGGCGCTTAAATAGAGCAATGACCGATC 2354  
 Db 301 GCATCTTCAATTTAGGAACTTGAATTAAGGCGCTTAAATAGAGCAATGACCGATC 360  
 Qy 2355 CATTTAAAGGTGATTTAAATTTTCTTCTTATCTTTTGGTGGCTTATGTTTATATAG 2414  
 Db 361 CATTTAAAGGTGATTTAAATTTTCTTCTTATCTTTTGGTGGCTTATGTTTATATAG 420  
 Qy 2415 TGTGGGTGGTGAATGATGTTCTTACACACTACCACTACCACTATATATATGGAA 2474  
 Db 421 TGTGGGTGGTGAATGATGTTCTTACACACTACCACTACCACTATATATATGGAA 480  
 Qy 2475 AATAGCTTCACTGCTAGA 2493  
 Db 481 AATAGCTTCACTGCTAGA 499

```

RESULT 4
AX111990
LOCUS AX111990 1708 bp DNA linear PAT 01-MAY-2001
DEFINITION Sequence 3 from Patent WO0125439.
ACCESSION AX111990
VERSION AX111990.1 GI:13938898
KEYWORDS
SOURCE
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
1
AUTHORS Bonello, J.F., Rogowsky, P. and Perez, P.
TITLE Plant seed endosperm-specific promoter
JOURNAL Patent: WO 0125439-A 3 12-APR-2001;
Biogenma (FR)
FEATURES
source
1..1708
Location/Qualifiers
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 13.7%; Score 341; DB 6; Length 1708;
Best Local Similarity 83.9%; Pred. No. 4.2e-61;
Matches 434; Conservative 0; Mismatches 75; Indels 8; Gaps 4;
Qy 1985 ATTCTCTCTTTGTCACCTTCGATATGACTTAATTTTATGTAGCAATGCCAATT 2044
Db 1192 ATTAAACATTTTGTGTCACCTTACAAAATGGCTTAATCTTTATGCGCAATGCCAATT 1251
Qy 2045 TTTTGTAGTGTGCCCTTATAGCACAAAACCTATATCCATTTCTTAATAGT-CCTTGA 2103
Db 1252 TTTTGTAGCGGTGCCCTTGTAAACATGAACTATTAATTTTCAATAGTACTTGA 1311
Qy 2104 AATCCACATTTATTTTGTAGCCATTTTCAAAATTTGGCACA----AACTAGGAAATTT 2159
Db 1312 AATTCGCATTTATTTTATGCAATTTTCAAAATTTGGACAAATTAATAGGAAATTC 1371
Qy 2160 AATACATCTTCCCAATACATATTTAGTGCATTTAGCAATTTAGTATGCTCAATATAGC 2219
Db 1372 AATACATCTTCCCAATACATATTTAGTGCATTTAGCAATTTAGTATGCTCAATATAGC 1431
Qy 2220 AACTCTCTTTGTAGATTCATTAATATGCTACATTTGCTACTT- -TTTGTAGATTTCA 2277
Db 1432 ACACATCTTTTGGACGATTAATAGTATTTGCTCACTACATTTCTTTGTAGCAGTTCA 1491
Qy 2278 TCAATAATGCTCATTTAGCATCTTCATTTTGGCAATTTGATTAACCGCTTAAATA 2337
Db 1492 TCAATAATGCTCATTTAGCATCTTCATTTTGGCAATTTGATTAACCGCTTAAATA 1551
Qy 2338 GAGCCAAAGTACCGATCCA- -TTTAAAGGTGATTTCTTAATTTCTTACTTCTTCTTGT 2396
Db 1552 GGGCCAAAGTACCAATCCACTTCNAAGGTGACTCTCTATTTCTTACTTCTTCTTGT 1611
Qy 2397 GGCTATGTTTATATATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2456
Db 1612 TGTATTTGTATATATATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1671
Qy 2457 TTGACATATATATGGAATATAGCTTTCACAGTCTAGA 2493
Db 1672 TTAGACATATATGGAATATAGCTTTCACAGTCTAGA 1708

RESULT 5
ZMA251319
LOCUS ZMA251319 1708 bp DNA linear PLN 26-APR-2000
DEFINITION Zea mays Ear3 gene, 5' flanking region.
ACCESSION AJ251319
VERSION AJ251319.1 GI:7657938

```

```

KEYWORDS 5' flanking region; Ear3 gene.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE
1
AUTHORS Bonello, J.F., Opsahl-Perstad, H.G., Perez, P., Dumas, C. and
Rogowsky, P.M.
TITLE Ear genes show different levels of expression in the same region of
maize endosperm
JOURNAL Gene 246 (1-2), 219-227 (2000)
MEDLINE 20331747
PUBMED 10767543
REFERENCE
2 (bases 1 to 1708)
AUTHORS Rogowsky, P.M.
TITLE Direct Submission
JOURNAL Submitted (24-NOV-1999) Rogowsky P.M., Reproduction et
Developpement des Plantes, ENS-Lyon, 46 allée d'Italie, F-69364
Lyon Cedex 07, FRANCE
FEATURES
Location/Qualifiers
source
1..1708
/organism="Zea mays"
/mol_type="genomic DNA"
/variant="hybrid DH5αDH7"
/db_xref="taxon:4577"
/map="chromosome arm 1S"
/tissue_type="leaf"
/dev_stage="seedling"
misc_feature
<1..276
/note="homology to non-LTR parts of several maize
retrotransposons"
<277..936
LTR
/note="homology to several maize retrotransposons"
1206..1708
gene
/gene="Ear3"
1306..1708
promoter
/gene="Ear3"
/note="homology to Ear1 and Ear2 promoters"
ORIGIN
Query Match 13.7%; Score 341; DB 8; Length 1708;
Best Local Similarity 83.9%; Pred. No. 4.2e-61;
Matches 434; Conservative 0; Mismatches 75; Indels 8; Gaps 4;
Qy 1985 ATTCTCTCTTTGTCACCTTCGATATGACTTAATTTTATGTAGCAATGCCAATT 2044
Db 1192 ATTAAACATTTTGTGTCACCTTACAAAATGGCTTAATCTTTATGCGCAATGCCAATT 1251
Qy 2045 TTTTGTAGTGTGCCCTTATAGCACAAAACCTATATCCATTTCTTAATAGT-CCTTGA 2103
Db 1252 TTTTGTAGCGGTGCCCTTGTAAACATGAACTATTAATTTTCAATAGTACTTGA 1311
Qy 2104 AATCCACATTTATTTTGTAGCCATTTTCAAAATTTGGCACA----AACTAGGAAATTT 2159
Db 1312 AATTCGCATTTATTTTATGCAATTTTCAAAATTTGGACAAATTAATAGGAAATTC 1371
Qy 2160 AATACATCTTCCCAATACATATTTAGTGCATTTAGCAATTTAGTATGCTCAATATAGC 2219
Db 1372 AATACATCTTCCCAATACATATTTAGTGCATTTAGCAATTTAGTATGCTCAATATAGC 1431
Qy 2220 AACTCTCTTTGTAGATTCATTAATATGCTACATTTGCTACTT- -TTTGTAGATTTCA 2277
Db 1432 ACACATCTTTTGGACGATTAATAGTATTTGCTCACTACATTTCTTTGTAGCAGTTCA 1491
Qy 2278 TCAATAATGCTCATTTAGCATCTTCATTTTGGCAATTTGATTAACCGCTTAAATA 2337
Db 1492 TCAATAATGCTCATTTAGCATCTTCATTTTGGCAATTTGATTAACCGCTTAAATA 1551
Qy 2338 GAGCCAAAGTACCGATCCA- -TTTAAAGGTGATTTCTTAATTTCTTACTTCTTCTTGT 2396
Db 1552 GGGCCAAAGTACCAATCCACTTCNAAGGTGACTCTCTATTTCTTACTTCTTCTTGT 1611
Qy 2397 GGCTATGTTTATATATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2456
Db 1612 TGTATTTGTATATATATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1671
Qy 2457 TTGACATATATATGGAATATAGCTTTCACAGTCTAGA 2493
Db 1672 TTAGACATATATGGAATATAGCTTTCACAGTCTAGA 1708

```







```

QY 2412 ATGTGCGGTGTTGAATGATGTTCTTACACCACTACACCAACACGTTGGACATATATATG 2471
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 ATATGTTTCATGTTGAGTGTATGTTCTTACACCACTACACCAACACGTTAGATATATATACA 509
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2472 GAAATAGCTTCACAGTCTAGA 2493
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 510 GAAATAGCTTCACATCTAGA 531
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
AF546188 203363 bp DNA linear PLN 20-DEC-2002
LOCUS Contiguous genomic DNA sequence comprising the 19-kDa zein gene
DEFINITION family from Zea mays, complete sequence.
ACCESSION AF546188
VERSION AF546188.1 GI:25992761
KEYWORDS HTG.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Song R. and Messing, J.
TITLE Contiguous Genomic DNA Sequence Comprising the 19-kD Zein Gene
JOURNAL Plant Physiol. 130 (4), 1626-1635 (2002)
PUBMED 12481046
REFERENCE 2 (bases 1 to 203363)
AUTHORS Song, R. and Messing, J.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2002) The Plant Genome Initiative at Rutgers -
Wakman Institute, Rutgers University, 190 Frelinghuysen Road,
Piscataway, NJ 08854-8020, USA
COMMENT Zea mays (B73) BAC clone Z492M16 containing 19 kDa zein z1B
subfamily, genomic location 1.
FEATURES
source
1. 203363
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMBB-492M16"
/clone_lib="CHORI201 library Mbol"

ORIGIN
Query Match 3.1%; Score 77.6; DB 8; Length 203363;
Best Local similarity 76.6%; Pred. No. 7.7e-06;
Matches 95; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 AGCTTTTCCGGTGAAGACGCTGTTAATCTTACAGCATGCTGGAACCAATAGTTA 60
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6526 AGCGTTTTCGTGACAAAGCTCTGTATATTATTAGTAGTAGCTTAAGCTTAAGCAATAGTTA 6585
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 GCTGTGTTTTGAGACCTTCGGAGAGTGAAGGCCCCCAACATCCCATCAATCAAGTC 120
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6586 GCATATTTTGGAGAACCTTCGGAGAGTGAAGGCCCCCAACATCAATCAATCAATCTT 6645
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 CCA 124
Db 6646 ACTA 6649

RESULT 12
AC139056 81372 bp DNA linear HTG 02-MAY-2003
LOCUS Mus musculus chromosome 15 clone RP23-17A19 map 15, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
ACCESSION AC139056
VERSION AC139056.2 GI:30315764
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Mus musculus (house mouse)

```

## ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 81372)

## AUTHORS

Birren, B., Nusbaum, C. and Lander, E.

## TITLE

Mus musculus chromosome 15, clone RP23-17A19

## JOURNAL

Unpublished

## REFERENCE

2 (bases 1 to 81372)

## AUTHORS

Anderson, S., Arachchi, H.M., Barna, N., Baatien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talanas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (23-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## REFERENCE

3 (bases 1 to 81372)

## AUTHORS

Anderson, S., Arachchi, H.M., Barna, N., Baatien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talanas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (02-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On May 2, 2003 this sequence version replaced gi:27877333. All repeats were identified using RepeatMasker:

Smit, A.P.A. &amp; Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submission@genome.wi.mit.edu

----- Project Information

Center project name: L29023

Center clone name: 17\_A\_19

-----

\* NOTE: This record contains 80 individual

\* sequencing reads that have not been assembled into

\* contigs. Runs of N are used to separate the reads

\* and the order in which they appear is completely

\* arbitrary. Low-pass sequence sampling is useful for

- \* Identifying clones that may be gene-rich and allows
- \* overlap relationships among clones to be deduced.
- \* However, it should not be assumed that this clone
- \* will be sequenced to completion. In the event that
- \* the record is updated, the accession number will
- \* be preserved.

```

1 906: contig of 906 bp in length
907 1006: gap of 100 bp
908 1006: contig of 906 bp in length
909 1006: contig of 906 bp in length
910 1006: gap of 100 bp
911 1006: contig of 906 bp in length
912 1006: contig of 906 bp in length
913 1006: gap of 100 bp
914 1006: contig of 906 bp in length
915 1006: contig of 906 bp in length
916 1006: gap of 100 bp
917 1006: contig of 906 bp in length
918 1006: contig of 906 bp in length
919 1006: gap of 100 bp
920 1006: contig of 906 bp in length
921 1006: contig of 906 bp in length
922 1006: gap of 100 bp
923 1006: contig of 906 bp in length
924 1006: contig of 906 bp in length
925 1006: gap of 100 bp
926 1006: contig of 906 bp in length
927 1006: contig of 906 bp in length
928 1006: gap of 100 bp
929 1006: contig of 906 bp in length
930 1006: contig of 906 bp in length
931 1006: gap of 100 bp
932 1006: contig of 906 bp in length
933 1006: contig of 906 bp in length
934 1006: gap of 100 bp
935 1006: contig of 906 bp in length
936 1006: contig of 906 bp in length
937 1006: gap of 100 bp
938 1006: contig of 906 bp in length
939 1006: contig of 906 bp in length
940 1006: gap of 100 bp
941 1006: contig of 906 bp in length
942 1006: contig of 906 bp in length
943 1006: gap of 100 bp
944 1006: contig of 906 bp in length
945 1006: contig of 906 bp in length
946 1006: gap of 100 bp
947 1006: contig of 906 bp in length
948 1006: contig of 906 bp in length
949 1006: gap of 100 bp
950 1006: contig of 906 bp in length
951 1006: contig of 906 bp in length
952 1006: gap of 100 bp
953 1006: contig of 906 bp in length
954 1006: contig of 906 bp in length
955 1006: gap of 100 bp
956 1006: contig of 906 bp in length
957 1006: contig of 906 bp in length
958 1006: gap of 100 bp
959 1006: contig of 906 bp in length
960 1006: contig of 906 bp in length
961 1006: gap of 100 bp
962 1006: contig of 906 bp in length
963 1006: contig of 906 bp in length
964 1006: gap of 100 bp
965 1006: contig of 906 bp in length
966 1006: contig of 906 bp in length
967 1006: gap of 100 bp
968 1006: contig of 906 bp in length
969 1006: contig of 906 bp in length
970 1006: gap of 100 bp
971 1006: contig of 906 bp in length
972 1006: contig of 906 bp in length
973 1006: gap of 100 bp
974 1006: contig of 906 bp in length
975 1006: contig of 906 bp in length
976 1006: gap of 100 bp
977 1006: contig of 906 bp in length
978 1006: contig of 906 bp in length
979 1006: gap of 100 bp
980 1006: contig of 906 bp in length
981 1006: contig of 906 bp in length
982 1006: gap of 100 bp
983 1006: contig of 906 bp in length
984 1006: contig of 906 bp in length
985 1006: gap of 100 bp
986 1006: contig of 906 bp in length
987 1006: contig of 906 bp in length
988 1006: gap of 100 bp
989 1006: contig of 906 bp in length
990 1006: contig of 906 bp in length
991 1006: gap of 100 bp
992 1006: contig of 906 bp in length
993 1006: contig of 906 bp in length
994 1006: gap of 100 bp
995 1006: contig of 906 bp in length
996 1006: contig of 906 bp in length
997 1006: gap of 100 bp
998 1006: contig of 906 bp in length
999 1006: contig of 906 bp in length
1000 1006: gap of 100 bp

```

```

34780: gap of 100 bp
34781: contig of 921 bp in length
34782: gap of 100 bp
34783: contig of 936 bp in length
34784: gap of 100 bp
34785: contig of 915 bp in length
34786: gap of 100 bp
34787: contig of 917 bp in length
34788: gap of 100 bp
34789: contig of 913 bp in length
34790: gap of 100 bp
34791: contig of 913 bp in length
34792: gap of 100 bp
34793: contig of 897 bp in length
34794: gap of 100 bp
34795: contig of 945 bp in length
34796: gap of 100 bp
34797: contig of 929 bp in length
34798: gap of 100 bp
34799: contig of 922 bp in length
34800: gap of 100 bp
34801: contig of 928 bp in length
34802: gap of 100 bp
34803: contig of 923 bp in length
34804: gap of 100 bp
34805: contig of 924 bp in length
34806: gap of 100 bp
34807: contig of 921 bp in length
34808: gap of 100 bp
34809: contig of 893 bp in length
34810: gap of 100 bp
34811: contig of 922 bp in length
34812: gap of 100 bp
34813: contig of 922 bp in length
34814: gap of 100 bp
34815: contig of 902 bp in length
34816: gap of 100 bp
34817: contig of 906 bp in length
34818: gap of 100 bp
34819: contig of 917 bp in length
34820: gap of 100 bp
34821: contig of 946 bp in length
34822: gap of 100 bp
34823: contig of 893 bp in length
34824: gap of 100 bp

```

Query Match 3.14; Score 76.4; DB 2; Length 81372;  
 Best Local Similarity 84.34; Pred. No. 1.4e-05;

Matches 86; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```

Qy 1 AAGCTTTTCGGTGTGATGAGCCTGTATATCTTAACAGCATCTCAAAACAAATAGTTA 60
    |||||
Db 70735 AAGCTTTTCGGTGTGAGCCTGTATATCTTAACAGCATCTCAAAACAAATAGTTA 70794

Qy 61 GCTGTGTTTTTGAGGACCTTCGGAAGATGAGGCCCCCAACA 102
    |||||
Db 70795 GCTGTGTTTTTGAGGACCTTCGGAAGATGAGGCCCCCAACA 70836

```

RESULT 13  
 AC145452/c

LOCUS AC145452 105035 bp DNA linear HTG 16-JUN-2004  
 DEFINITION Zea mays clone ZMMBSc0016A01, ... SEQUENCING IN PROGRESS ...

ACCESSION AC145452

VERSION AC145452.7 GI:48717633

KEYWORDS HTG; HTGS\_PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.

SOURCE Zea mays

ORGANISM Zea mays  
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Zea.

Center clone name: 16\_A.1  
----- Consensus Information  
This consensus is derived from a shotgun assembly that has been manually curated. It is the best representation of the BAC that we can generate without further laboratory work. The draft assembly has been edited, and if possible, ends identified by vector as well as by BAC end sequences, and contigs ordered and oriented. Bases that are not Ns are either above Q20 or manually edited. This assembly was performed with Arachne (Genome Res. 2002 12: 177-189; Genome Res. 2003 13: 91-96). All trace files for this project are available at the NCBI trace repository (http://www.ncbi.nlm.nih.gov/trace/trace.fcgi?). An exact list of reads used in this assembly are available at http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.  
-----  
\* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 105035: contig of 105035 bp in length.  
Location/Qualifiers  
1. 105035  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMWBC0016A01"  
/clone\_lib="CHORI-201 Maize B73 BAC EScori Library"  
  
FEATURES  
Source  
105035  
Query Match 3.1%; Score 76.4; DB 2; Length 105035;  
Best Local Similarity 84.3%; Pred. No. 1.4e-05;  
Matches 86; Conservative 0; Mismatches 16; Indels 0; Gaps 0;  
  
ORIGIN  
1 AACCTTTTCCGTCATGACGACCTCTGTAATCTTAAACAGCATCTGAAACAAATAGTTA 60  
Db 103873 AGCTCTTCTGTCGACGAGCTCTGTAGTATTTCAGTATATGCTGAAACAAATAGTTA 103814  
  
Qy 61 GCTGTGTTTTGACGACCTCTGAGATGAGGCCCCCA 102  
Db 103813 GCTGTGTTTTGACGACCTCTGAGATGAGGCCCTCCCA 103772  
  
RESULT 14  
166494/c  
LOCUS 166494  
DEFINITION Sequence 14 from patent US 5670367.  
ACCESSION 166494  
VERSION 166494.1 GI:2724471  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 7218)  
AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.  
TITLE Recombinant fowlpox virus  
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;  
FEATURES  
1. 7218  
Location/Qualifiers  
/organism="unknown"  
/mol\_type="unassigned DNA"  
  
ORIGIN  
  
Query Match 3.0%; Score 74; DB 6; Length 7218;  
Best Local Similarity 9.0%; Pred. No. 4.9e-05;  
Matches 41; Conservative 234; Mismatches 179; Indels 0; Gaps 0;  
  
Qy 1286 AGTGAGGGTGAATGTCAGTTCTGTCACGTCGGGAATAGTGGCGCATCGCTGTGAATGAAT 1345



/cultivar="B73"  
/db\_xref="taxon:4577"  
/clone="ZMHB0375A01"  
/clone\_lib="CHORI-201 Maize B73 BAC MboI Library"

ORIGIN

Query Match 2.9%; Score 72.6; DB 2; Length 192816;  
Best Local Similarity 67.5%; Pred. No. 8.7e-05;  
Matches 102; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 AAGCTTTCCGGTGAAGCAACCTGTAATCTTAACAGCATGCTGGAACAAATAGTTA 60  
Db 142127 AACCTTTGTTATGACGAGCTCCTGTAATATTGACATTATGCTAGAAATTAATGGTTA 142068

QY 61 GCTGTGTTTTTGAGGACCTTCGGAAGATGAAGGCCCAACACATCCCATGCAATCAAGTC 120  
Db 142067 GTCGTGTTTTTGAGGACCTTCGGAACAGAGAGGCCCAACACACTCAATTGTATATGC 142008

QY 121 CCCATGACTTGCMAAAAGCAATTTTATCA 151  
Db 142007 ATTATCATGAACATTCGAAGAACTTTATTA 141977

Search completed: April 9, 2005, 05:51:59  
Job time : 10363 secs